

GOWERS NORMS CONTROL DIOPHANTINE INEQUALITIES

ALED WALKER

ABSTRACT. The classical ‘Generalised von Neumann Theorem’ is a central tool in the study of systems of linear equations with integer coefficients. This theorem reduces the task of counting weighted solutions of these equations to that of counting the weighted solutions for a particular family of forms, the Gowers norms $\|f\|_{U^{s+1}[N]}$ of the weight f . In this paper we consider the related study of systems of linear inequalities with real coefficients, and show that the number of solutions to such weighted Diophantine inequalities may also be bounded by Gowers norms, provided that the weights themselves are bounded. Further, we provided a necessary and sufficient condition for a system of real linear forms to be governed by Gowers norms in this way. In a future paper we will discuss the case in which the weights are unbounded but suitably pseudorandom, with applications to solving Diophantine inequalities over the primes.

1. INTRODUCTION

Diophantine inequalities are a vast and varied topic in analytic number theory (see [2], say). We will focus on a particular class of problems, which may all be represented in the following general form. Let $A \subset \mathbb{Z}$ be some fixed subset of integers, $\varepsilon > 0$, and $L = (\lambda_{ij})$ some m -by- d real matrix. Viewing A^d as a subset of \mathbb{R}^d , one may ask whether there are infinitely many solutions to

$$\|L\mathbf{a}\|_\infty \leq \varepsilon \tag{1}$$

with $\mathbf{a} \in A^d$. Further, one might seek an asymptotic formula for the number of solutions with $\mathbf{a} \in A^d \cap [1, N]^d$ as N tends to infinity. Much is known about this problem for certain special sets A – most notably the image sets of polynomials – although, as far as the author is aware, the situation has not before been considered in full generality.

We will state our main theorem at the end of this introduction, as the full statement is moderately technical. However, in summary, we successfully show that the (normalised) number of solutions to (1) with $\mathbf{a} \in A^d \cap [1, N]^d$ is bounded by a function of a Gowers norm¹ $\|1_A\|_{U^{s+1}[N]}$, provided that L satisfies a particular set of non-degeneracy conditions. As an immediate application, we will find asymptotic expressions for the number of solutions to (1) when A is a positive density pseudorandom set, and, by using a deep result of [11], we will be able to count solutions weighted by the Möbius function μ . All these results will, generically, only require $d \geq m + 2$, and are thus much stronger than those which may be achieved by classical analytic methods alone, which typically require $d \geq 2m + 1$.

Additionally to the above, we will provide examples to demonstrate that, for Gowers norms to be applied in this way, our non-degeneracy condition on L is necessary. The analysis of this condition – which may be viewed as an approximate real-variable version of that which occurs in [10, Theorem 1.8] – and the process by which we transfer from functions on \mathbb{Z} to functions on \mathbb{R} together constitute the main new work of this paper. In general our arguments follow [10] very closely, though the reader may find useful our exposition of a generalised von Neumann theorem purely for bounded functions, as the

¹The basic theory of Gowers norms necessary for a comprehension of the proof will be developed in Appendix A.

relevant appendices of [10] describe the theory in the full generality of functions bounded by a pseudorandom measure.

It has already been noted that much is known about our problem for certain special sets A , particularly when $m = 1$, and we take the opportunity to recall some of the main classical results in the area. If A is the set of squares, say, it was shown by Davenport and Heilbronn in [4] for $m = 1$ and $d = 5$ (provided the coefficients λ_i are non-zero, not all of the same sign, and not all in pairwise rational ratio) that there are infinitely many solutions to (1), i.e. infinitely many solutions to

$$|\lambda_1 n_1^2 + \lambda_2 n_2^2 + \lambda_3 n_3^2 + \lambda_4 n_4^2 + \lambda_5 n_5^2| \leq \varepsilon.$$

Their work also proves the same result for k^{th} powers, provided that the number of variables is at least $2^k + 1$. The method is Fourier analytic – replacing the interval $[-\varepsilon, \varepsilon]$ with a smooth cut-off and expressing the solution count via the inversion formula. See [3, Chapter 20], [23, Chapter 11]. Freeman [6] refined the minor-arc analysis from [4] to obtain asymptotic formulas for the number of solutions with $n_i \leq N$ for all i . Of course there is much more work on such polynomial questions, only tangentially related to this paper, i.e. Margulis’ solution to the Oppenheim Conjecture [15]. Regarding questions with $m \geq 2$, [19] considers once again the case of A being the k^{th} powers, [17] developing a refined result in the case of inequalities for general real quadratics.

These questions have also been asked when A is the set of prime numbers, and may be tackled using similar analytic tools. A result first claimed in [1] by Baker² states that for any fixed $\varepsilon > 0$ there exist infinitely many triples of primes (p_1, p_2, p_3) satisfying

$$|\lambda_1 p_1 + \lambda_2 p_2 + \lambda_3 p_3| \leq \varepsilon, \tag{2}$$

assuming again that the coefficients λ_i are non-zero, not all of the same sign, and not all in pairwise rational ratio. Parsell [18] then used a similar refinement to that of Freeman to prove a lower bound³ on the number of solutions to (2) of the expected order of magnitude $\asymp \varepsilon N^2 (\log N)^{-3}$.

These analytic approaches ultimately rely on establishing tight mean-value estimates for certain exponential sums, and thus require a large enough number of variables for such estimates to hold. In the case of primes, say, for m inequalities the method of Parsell will yield an asymptotic for the number of solutions to (1) in prime variables provided $d \geq 2m + 1$, at least for generic L .

It is natural to view these questions regarding linear inequalities as closely related to the analogous questions involving linear equations with rational coefficients⁴. The study of such systems was greatly enhanced by the introduction, by Green and Tao in [10], of a powerful and wide-ranging technique, known as a ‘Generalised von Neumann Theorem’, which can be used to show that certain special systems of linear equations – known as Gowers norms – are, in some sense, ‘universal’ over all linear systems; this is [10, Theorem 7.1]. It was using this technique, in combination with a deep study of the inverse theory of these Gowers norms, that those authors and Ziegler managed to prove that, generically, $m + 2$ prime variables are adequate to obtain an asymptotic formula for the number of prime solutions to m linear equations with rational coefficients, rather

²In fact Baker proved a slightly different result, writing in the cited paper that the result we quote here followed easily from the then existing methods. A proof does not seem to have been written down until Parsell [18].

³An asymptotic formula for the number of solutions follows very easily from Parsell’s work, though does not appear to be present in the literature.

⁴Indeed, if all the coefficients of L are rational the two problems are essentially identical, a fact which will be used in our main proof.

than the $2m + 1$ variables required by the circle method.

Though ultimately we wish to understand the case where A is the set of primes, the purpose of this article is rather to develop a theory for very general positive density sets A which will strictly generalise the theory available in the case of linear Diophantine equations with rational coefficients. This will be achieved by adapting linear algebra arguments over \mathbb{Q} and $\mathbb{Z}/N\mathbb{Z}$ to go through over \mathbb{R} . As in [10] we will work at a fixed large scale N , in particular not assuming that the coefficients λ_{ij} of L are independent of N (although they will be $O(1)$). This gives our main theorem the widest possible application, yet also introduces new technical problems, not merely with finding quantitative statements of various linear algebra lemmas but also with issues of rational approximation.

In the main body of the paper we will assume an easy familiarity with the definitions of Gowers norms over \mathbb{Z} , and to some extent over \mathbb{R} . However, for the non-expert, the necessary theory is described in Appendix A.

Let us introduce a multi-linear form which will count solutions to (1) weighted by general functions f_i . Let $f_1, \dots, f_d : [N] \rightarrow [-1, 1]$. Then define

$$T_\varepsilon(f_1, \dots, f_d) := \frac{1}{N^{d-m}} \sum_{\substack{\mathbf{n} \in [N]^d \\ L\mathbf{n} \in B_\varepsilon^m}} \prod_{i=1}^d f_i(n_i),$$

where B_ε^m denotes the closed ball in \mathbb{R}^m with radius ε in the L^∞ norm.

This definition deserves a brief comment. One could argue that it is more natural to normalise the above summation not by N^{-d+m} but rather by the count $|\{\mathbf{n} \in [N] : L\mathbf{n} \in B_\varepsilon^m\}|^{-1}$ itself, which in applications is a quantity of size approximately $\varepsilon^{-a} N^{-d+m}$, for some $a \in [m]$. The value of a depends on diophantine properties of L , in particular whether there is some vector $\alpha \in \mathbb{R}^m$ such that $\alpha^T L \in \mathbb{Z}^d$. We prefer to avoid the existence of uncertain ε dependence in a normalising factor, but the dependence is tame, and one could set up an identical theory with the alternative normalisation if one so wished. The key point is that, under the hypotheses of the main theorem, $T_\varepsilon(f_1, \dots, f_d)$ should be bounded. This is proved in Proposition 12.

We say that L satisfies the *von Neumann property* if for arbitrary functions $f_1, \dots, f_d : [N] \rightarrow [-1, 1]$,

$$|T_\varepsilon(f_1, \dots, f_d)| \ll \kappa(\rho) + o_\rho(1), \quad (3)$$

for some⁵ $k \leq d - 1$, provided that $\|f_i\|_{U^k[N]} \leq \rho$ for some f_i . The implied constant is independent of N , and the choice of f_i . Here we have used κ notation: a function $F : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ satisfies $F(x) = \kappa(x)$ if $F(x) \rightarrow 0$ as $x \rightarrow 0$.

Before we state our main theorem, we also need to introduce a notion to help formulate the non-degeneracy condition on L . We say that an m -by- d real matrix L' , with $d \geq m + 2$, is ‘in the dual degeneracy variety’ if there is a non-zero vector in the row-space of L with two or fewer non-zero entries. We denote this variety by V_{degen}^* . The relevant notion of non-degeneracy for us will be that the matrix L should be bounded away from the dual degeneracy variety⁶.

⁵The monotonicity property of Gowers norms means that we may assume without loss of generality that $k = d - 1$, but it does not seem necessary to do this.

⁶There is a related condition required for [10, Theorem 1.8], in which, stated in this language, one merely requires the matrix L from (1) to be in the complement of V_{degen}^* .

Finally define

$$\text{dist}(L, V) := \inf_{L' \in V} \|L - L'\|_\infty,$$

where the L^∞ norm is taken on \mathbb{R}^{md} .

We may now state the main theorem of this paper, which states that, provided L doesn't come close to satisfying a certain 'measure zero' algebraic condition, Gowers norms may be used to bound the weighted solution count.

Theorem 1 (Main Theorem). *Let $\varepsilon, c, C > 0$ be fixed. Let $L = L(N)$ be an m -by- d real matrix of rank m , with $d \geq m + 2$, and with $\|L\|_\infty \leq C$. Suppose further that $\|L - L'\|_\infty \geq c$ for all m -by- d matrices L' with $\text{rank}(L') < m$. Then, if $\text{dist}(L, V_{\text{degen}}^*) \geq c'$ for some absolute $c' > 0$, L satisfies the von Neumann property (with implied constant depending on ε, c, c' , and C).*

We note immediately that we may in fact replace c and c' by a single constant c in the above statement without weakening the conclusion, so we proceed with this assumption.

Let us illustrate Theorem 1 with three key examples.

Example 2 (Three-term irrational AP). The first example may be proved using the classical Davenport-Heilbronn method, but we include it to demonstrate the simplest case where Theorem 1 applies. If

$$L = \begin{pmatrix} 1 & -\sqrt{2} & -1 + \sqrt{2} \end{pmatrix},$$

then $m = 1$ and $d = 3$, and the sizes of the coefficients of L are bounded by $\sqrt{2}$. For rather trivial reasons, $\|L - L'\|_\infty \geq \sqrt{2}$ for all 1-by-3 matrices L' with $\text{rank}(L') < 1$, i.e. for the matrix $L' = 0$. Further, $\text{dist}(L, V_{\text{degen}}^*) \geq -1 + \sqrt{2}$, as one must change some co-ordinate of L by at least $-1 + \sqrt{2}$ in order to reach a matrix with only two non-zero entries. Therefore Theorem 1 applies, and we conclude that

$$\frac{1}{N^2} \left| \sum_{\substack{n_1, n_2, n_3 \leq N \\ |n_1 - \sqrt{2}n_2 + (-1 + \sqrt{2})n_3| \leq \varepsilon}} f_1(n_1)f_2(n_2)f_3(n_3) \right| \ll_\varepsilon \min \kappa(\rho) + o_\rho(1),$$

for any three bounded functions f_i such that $\min \|f_i\|_{U^2[N]} \leq \rho$.

This may be interpreted in terms of counting the number of a certain irrational pattern, a 'three-term irrational arithmetic progression'. Indeed, for $\theta \in \mathbb{R}$ let $\langle \theta \rangle$ denote $\lfloor \theta + \frac{1}{2} \rfloor$, i.e. the nearest integer to θ . Then for any three bounded functions $f_1, f_2, f_3 : [N] \rightarrow \mathbb{R}$ consider

$$T = \frac{1}{N^2} \sum_{x, d \in \mathbb{Z}} f_3(x)f_2(x+d)f_1(\langle x + \sqrt{2}d \rangle).$$

T counts the number of near-occurrences of the pattern $(x, x+d, x+\sqrt{2}d)$, weighted by the f_i .

By a simple change of variables $n_1 = \langle x + \sqrt{2}d \rangle$, $n_2 = x+d$, $n_3 = x$, and noting that $x + \sqrt{2}d \notin \frac{1}{2}\mathbb{Z}$, we see

$$T = \sum_{\substack{n_1, n_2, n_3 \leq N \\ |n_1 - \sqrt{2}n_2 + (-1 + \sqrt{2})n_3| \leq \frac{1}{2}}} f_1(n_1)f_2(n_2)f_3(n_3).$$

By the above, we have shown that this expression is $O(\kappa(\rho) + o_\rho(1))$, provided $\min \|f_i\|_{U^2[N]} \leq \rho$.

Now, suppose $|A| = \alpha N$ for some $\alpha > 0$, and let $b := 1_A - \alpha 1_{[N]}$ be the so-called ‘balanced function’ of A . By the usual telescoping trick, $T_{\frac{1}{2}}(1_A, 1_A, 1_A)$ is equal to

$$T_{\frac{1}{2}}(\alpha 1_{[N]}, \alpha 1_{[N]}, \alpha 1_{[N]}) + T_{\frac{1}{2}}(b, \alpha 1_{[N]}, \alpha 1_{[N]}) + T_{\frac{1}{2}}(1_A, b, \alpha 1_{[N]}) + T_{\frac{1}{2}}(1_A, 1_A, b).$$

Bounding each term on the right in terms of $\|b\|_{U^2[N]}$, one may establish that

$$\frac{1}{N^2} \sum_{x, d \in \mathbb{Z}} 1_A(x) 1_A(x+d) 1_A(\langle x + \sqrt{2}d \rangle) = C_1 \alpha^3 + \kappa(\rho) + o_\rho(1)$$

for some constant C_1 , provided $\|b\|_{U^2[N]} \leq \rho$.

Example 3 (Four-term irrational APs). If

$$L = \begin{pmatrix} 1 & 0 & -\sqrt{2} & -1 + \sqrt{2} \\ 0 & 1 & -\sqrt{3} & -1 + \sqrt{3} \end{pmatrix},$$

then $T_{\frac{1}{2}}(f_1, \dots, f_4)$ counts the number of patterns $(x, x+d, \langle x + \sqrt{2}d \rangle, \langle x + \sqrt{3}d \rangle)$, weighted by f_3, f_4, f_1 and f_2 respectively. This can be thought of as a four-term irrational arithmetic progression. Note that $\|L\|_\infty \leq \sqrt{3}$ and $\|L - L'\|_\infty \geq 1$ for all rank 1 matrices L' .

Now, when $d = m + 2$, elementary linear algebra shows that the condition $L \in V_{\text{degen}}^*$ is equivalent to the existence of some m -by- m submatrix with determinant zero. With L as above, we see that none of the 6 determinants of the 2-by-2 submatrices are zero, and so we conclude that $L \notin V_{\text{degen}}^*$. Since L is independent of N , there exists some constant $c > 0$ such that $\text{dist}(L, V_{\text{degen}}^*) \geq c$. Therefore Theorem 1 applies and, telescoping as above, we may derive

$$\frac{1}{N^2} \sum_{x, d \in \mathbb{Z}} 1_A(x) 1_A(x+d) 1_A(\langle x + \sqrt{2}d \rangle) 1_A(\langle x + \sqrt{3}d \rangle) = C \alpha^4 + \kappa(\rho) + o_\rho(1)$$

for some constant C , provided $\|b\|_{U^3[N]} \leq \rho$.

We comment that the infinitary theory of these particular patterns was previously considered in [16], albeit in the different language of ergodic theory. In particular, an easy deduction from [16, Theorem B] shows that all sets $A \subset \mathbb{N}$ with positive upper Banach density contain infinitely many copies of the pattern $(x, x+d, \langle x + \sqrt{2}d \rangle, \langle x + \sqrt{3}d \rangle)$.

Example 4. Theorem 1 has immediate consequences for counting solutions to Diophantine inequalities weighted by explicit bounded pseudorandom functions. In particular there is the following natural analogue of [10, Proposition 9.1].

Let L satisfy the hypotheses of Theorem 1, and μ be the Möbius function. Then $T_\varepsilon(\mu, f_2, \dots, f_d) = o_{C, c, c', \varepsilon}(1)$ for any bounded functions f_2, \dots, f_d . The same is true with μ replaced by the Liouville function λ . This follows immediately from Theorem 1 and the deep facts (stated in [10], proved in [11] and [14]) that $\|\mu\|_{U^k[N]} = o_k(1)$ and $\|\lambda\|_{U^k[N]} = o_k(1)$.

For example, this shows that if $\varepsilon < \frac{1}{2}$ then

$$\sum_{\substack{\mathbf{n} \in [N]^4 \\ n_1 - n_2 = n_2 - n_3 \\ |(n_2 - n_3) - \sqrt{2}(n_3 - n_4)| \leq \varepsilon}} \mu(n_1) \mu(n_2) \mu(n_3) \mu(n_4) = o_\varepsilon(N^2).$$

The left-hand side may be considered as counting another kind of four-term irrational arithmetic progression.

The proof of Theorem 1 is relatively straightforward for the previous examples, as L is constant in terms of N and the rows are linearly independent modulo \mathbb{Z} . The next example illustrates some of the difficulty in proving Theorem 1 in full generality.

Example 5. If

$$L = \begin{pmatrix} 1 + N^{-1} & \sqrt{3} + (\log N)^{-1} & \pi & -\pi + \sqrt{2} \\ 2 & 2\sqrt{3} + N^{\frac{1}{2}} & -\sqrt{5} & e \end{pmatrix},$$

then L has rank 2 and $L \notin V_{\text{degen}}^*$. One might therefore hope to apply the theory of Gowers norms to bound the number of solutions to inequalities given by L . However, by considering perturbations of the first two columns, we see that $\text{dist}(L, V_{\text{degen}}^*) = o(1)$. (Indeed, one may perturb L by $O((\log N)^{-1})$ such that there is a vector $(0, 0, x_3, x_4)$ in the row space). Therefore Theorem 1 does not apply in this case, despite the fact that $L \notin V_{\text{degen}}^*$.

In fact, the conclusion of Theorem 1 cannot possibly hold in the above example. Indeed, we have the following partial converse.

Theorem 6. *Let $\varepsilon, c, C > 0$ be fixed. Let $L = L(N)$ be a family of m -by- d matrices of rank m , with $d \geq m + 2$, and with $\|L\|_\infty \leq C$. Suppose further that $\|L - L'\|_\infty \geq c$ for all m -by- d matrices L' with $\text{rank}(L') < m$, and that $T_\varepsilon(1, \dots, 1) \gg_{c, C, \varepsilon} 1$ for large enough N . Then, if*

$$\liminf_{N \rightarrow \infty} \text{dist}(L, V_{\text{degen}}^*) = 0,$$

L does not satisfy the von Neumann property (for any $k \in \mathbb{N}$, and for any value of the implied constant).

In the other words, for any choice of implied constant K , $\kappa(x)$ function, $k \in \mathbb{N}$ and $o_\rho(1)$ error term, we can always find a large N and bounded functions f_1, \dots, f_d such that for some $i \in [d]$ we have $\|f_i\|_{U^k[N]} > \rho$ but

$$|T_\varepsilon(f_1, \dots, f_d)| > K\kappa(\rho) + o_\rho(1).$$

We now briefly discuss the method of proof for Theorem 1. It is tempting to think that the theorem will follow easily from taking rational approximations of the coefficients λ_{ij} , and then using the existing Generalised von Neumann Theorem over \mathbb{Q} as a black box. Though of course we do not rule out an alternative approach to that of this paper, it seems that such an argument will only quickly succeed if the coefficients λ_{ij} are all extremely well-approximable – else the height of the rational approximations becomes too great to apply the classical Generalised von Neumann Theorem – and that one must find an alternative method for other L .

In section 3 we set out such a programme, and, in particular, we will deduce this Theorem 1 from a real-variable result, which may be thought of as a more usual Generalised von Neumann Theorem, but over \mathbb{R} . For functions $g_1, \dots, g_d : [0, N] \rightarrow [-1, 1]$, define

$$\tilde{T}_\varepsilon(g_1, \dots, g_d) := \frac{1}{N^{d-m}} \int_{\mathbf{Lx} \in B_\varepsilon^m} \prod_{i=1}^d g_i(x_i) d\mathbf{x},$$

and say that L satisfies the *real von Neumann property* if for arbitrary $g_1, \dots, g_d : [0, N] \rightarrow [-1, 1]$, we have

$$\left| \tilde{T}_\varepsilon(g_1, \dots, g_d) \right| \ll \kappa(\rho) + o_\rho(1). \quad (4)$$

for some $k \leq d - 1$, provided that $\|g_i\|_{U^k(\mathbb{R})} \leq \rho$ for some g_i .

Theorem 7 (Generalised von Neumann Theorem over \mathbb{R}). *Let $\varepsilon, c, C > 0$ be fixed. Let $L = L(N)$ be an m -by- d real matrix of rank m , with $d \geq m + 2$, and with $\|L\|_\infty \leq C$. Suppose further that $\|L - L'\|_\infty \geq c$ for all m -by- d matrices L' with $\text{rank}(L') < m$. Then, if $\text{dist}(L, V_{\text{degen}}^*) \geq c$, L satisfies the real von Neumann property (with implied constant depending on ε, c and C).*

In order to be able to deal with every possible matrix L , in particular those which exhibit hybrid behaviour between fully rational and fully equidistributing, we will need to formulate a suitable hybrid real-variable theorem. This technical refinement of Theorem 7 will be deferred to Theorem 25 in section 4.

The plan of the paper is as follows. In section 2 we will recall some linear algebra reductions from [10] and state the ‘quantitative normal form’ lemma we require. The proof of this is deferred to Appendix B as, though important to the thrust of the paper, the details reduce to standard arguments. In section 3 we will introduce the importance of the theory of equidistribution, and describe how Theorem 1 follows from Theorem 25 (the slight technical generalisation of Theorem 7). In section 4 we will marshal together all arguments and prove Theorem 25 (and hence Theorem 7) in full generality, hence completing the proof of Theorem 1. This is the section which most resembles Appendix C of [10], but certain key quantitative details will be different. In section 5 we will describe counterexample constructions of functions f_i which prove Theorem 6.

In the appendices we give an overview of the necessary theory of Gowers norms, and the details of the quantitative normal form argument. We also prove two easy propositions on quantitative equidistribution, and, in the final appendix, describe two properties of Lipschitz functions required in the main text.

Notation: We will use standard asymptotic notation O , o , and \ll throughout. The symbol $[N]$ will always denote $\{n \in \mathbb{N} : n \leq N\}$, whereas $[0, N]$ will be reserved for the closed real interval. For $x \in \mathbb{R}$, we write $\langle x \rangle := \lfloor x + \frac{1}{2} \rfloor$ for the nearest integer to x , and $\|x\|_{\mathbb{R}/\mathbb{Z}}$ or $\|x\|_{\mathbb{T}}$ for $|x - \langle x \rangle|$. Analogous notation will be used for higher dimensional tori. We will work on Euclidean space always with the $\|\cdot\|_\infty$ norm, and the operator dist will always be with respect to this norm. On a single occasion we will require the operator dist with respect to the $\|\cdot\|_2$ norm, where it will be denoted by dist_{Euc} .

As mentioned earlier, we also adopt the κ notation: $\kappa(\delta)$ is some quantity that tends to zero as δ tends to zero.

All implied constants will be allowed to depend on m and d . A lot of implied constants will depend on c and C too, but we choose to explicitly record when this is the case in order to highlight where these constants impact the proof. Consequently, the letters c and C will be reserved for the constants in the statement of Theorem 1, and so C_1 and c_1 will be used for other large and small constants respectively. In the customary fashion the exact values of c_1 and C_1 may vary from line to line. Occasionally we will need to pick additional constants which are small and large with respect to c_1 and C_1 already chosen: we will resort to c_2 and C_2 , and so on.

As a general rule we will not always write the formal $d\mathbf{x}$ expressions at the end of all integrals – as this would render several expressions unmanageably long – preferring to denote the integration as subscripts under the integral, analogous to notation for discrete summation. Integrals in which no limits are given are always assumed to be over all of \mathbb{R} .

Acknowledgements: The author would like to thank his doctoral supervisor Ben Green for his many suggestions and calm insight. A discussion with Alex Betts was also useful for clarifying certain arguments in Appendix C. Parts of the work were completed while the author was a Programme Associate at the Mathematical Sciences

Research Institute in Berkeley, which provided excellent working conditions. The author is supported by EPSRC grant no. EP/M50659X/1.

2. LINEAR ALGEBRA

Here we introduce some technical linear algebraic notions which will help to quantify the manipulations required in section 4, stating the main propositions we require. The proofs reduce to a close analysis of known algorithms, and would no doubt be obvious to the expert, and for these reasons we defer them to Appendix B.

Firstly, to prove Theorem 1 it will be useful to introduce of a concept we will refer to as the *rank matrix* of L . Consider the following basic fact: as L is assumed to have row-rank m , L also has column rank m , and thus we may find m linearly independent columns in L which determine an m -by- m submatrix M with non-zero determinant. When L depends on the quantity N , we use the term ‘rank matrix’ to refer to a quantitative refinement of the above. Indeed, we say M is a rank matrix of L if M is an m -by- m submatrix such that $\det M = \Omega(1)$ as N tends to infinity.

Proposition 8. *Suppose that $\|L\|_\infty \geq C$ and $\|L - L'\|_\infty \geq c$ for all matrices L' with $\text{rank}(L') < m$. Then L has a rank matrix M , with $|\det M| = \Omega_{c,C}(1)$. Furthermore, if \mathbf{v} is any vector in the row-space of L with coefficients bounded by C , then for $1 \leq i \leq m$ there exist coefficients a_i with $|a_i| = O_{c,C}(1)$ such that $\sum_{i=1}^m a_i \lambda_{ij} = v_j$ for all $1 \leq j \leq d$.*

We defer the proof to Appendix B. We leave here the main fact we will use about the rank matrix, which is that, if $\varepsilon = O_{c,C}(1)$, $M^{-1}B_\varepsilon^m \subseteq [-\frac{1}{10}, \frac{1}{10}]$. Perhaps the clearest proof of this is to note that the matrix $M^T M$ has coefficients bounded by $O_C(1)$, and hence all the eigenvalues of $M^T M$ are $O_C(1)$. This implies that all the eigenvalues of $M^{-1}(M^T)^{-1}$ are $\Omega_C(1)$. But $|\det(M^{-1}(M^T)^{-1})|$ is $O_{c,C}(1)$, by standard facts about the determinant and the assumption that M is a rank matrix. Then, considering the determinant as the product of eigenvalues and swapping the transposition and inversion operations, we combine these two observations to conclude that $\max\{|\lambda| : \lambda \text{ an eigenvalue of } M^{-1}(M^{-1})^T\}$ is $O_{c,C}(1)$. But then

$$\begin{aligned} \|M^{-1}\|_{op} &= \|M^{-1}(M^{-1})^T\|_{op}^{\frac{1}{2}} \\ &= (\max\{|\lambda| : \lambda \text{ an eigenvalue of } M^{-1}(M^{-1})^T\})^{\frac{1}{2}} \\ &= O_{c,C}(1), \end{aligned}$$

since $M^{-1}(M^{-1})^T$ is symmetric, so we may conclude that $M^{-1}B_\varepsilon^m \subseteq [-\frac{1}{10}, \frac{1}{10}]$ for small enough ε .

The second technical linear algebra condition required involves reparametrising certain linear forms in order to enable a particular application of the Cauchy-Schwarz inequality. Let $(\psi_1, \dots, \psi_m) = \Psi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a system of homogeneous linear forms. The crux of the theory from [10] is that, provided it is of finite Cauchy-Schwarz complexity, Ψ admits an extension which is in so-called normal form.

In words, a reparametrisation into normal form is one in which each linear form is the only one of forms which mentions all of its particular collection of variables. For example, the forms

$$\begin{aligned} \psi_1(t, u, v) &= u + v \\ \psi_2(t, u, v) &= v + t \\ \psi_3(t, u, v) &= u + t \\ \psi_4(t, u, v) &= u + v + t \end{aligned}$$

are in normal form with respect to ψ_4 , since it is the only form to utilise all three of the variables. However, this system is not in normal form with respect to ψ_3 , say. However, the system

$$\begin{aligned}\psi_1(t, u, v, w) &= u + v + 2w \\ \psi_2(t, u, v, w) &= v + t - w \\ \psi_3(t, u, v, w) &= u + t - w \\ \psi_4(t, u, v, w) &= u + v + t,\end{aligned}$$

which parametrises the same set of patterns, is in normal form for all i . We repeat the precise definition from [10].

Definition 9. Let $\Psi = (\psi_1, \dots, \psi_m)$ be a system of linear forms on \mathbb{R}^n , and let $s \geq 0$. We say that Ψ is in s -normal form with respect to ψ_i if there exists a collection $J_i \subseteq \{\mathbf{e}_1, \dots, \mathbf{e}_d\}$ of basis vectors of cardinality $|J_i| = s + 1$ such that

$$\prod_{\mathbf{e} \in J_i} \psi_{i'}(\mathbf{e})$$

is non-zero for $i' = i$ and vanishes otherwise. We say that Ψ is in s -normal form if it is in s -normal form with respect to ψ_i for every i .

Let us also recall what it means for a certain system of forms Ψ' to extend the systems of forms Ψ .

Definition 10. For a system of linear forms Ψ on \mathbb{R}^d , an extension Ψ' is a system of linear forms on $\mathbb{R}^{d'}$ for some $d' \geq d$ such that $\Psi'(\mathbb{R}^{d'}) = \Psi(\mathbb{R}^d)$ and furthermore if we identify \mathbb{R}^d with the subset $\mathbb{R}^d \times \{0\}^{d'-d}$ in the obvious manner then Ψ is the restriction of Ψ' to this subset.

The condition on L from Theorem 1 is introduced so that $\ker(L)$, a $d - m$ dimensional subspace of \mathbb{R}^d , may be parametrised

$$x_i = \psi_i(u_1, \dots, u_{d-m})$$

for $1 \leq i \leq d$, by system of forms Ψ , which, for each i , admits an extension Ψ' in normal form with respect to ψ'_i . Furthermore, we need Ψ' to be obtained from Ψ ‘by bounded shifts of the variables u_j ’, and also for all the non-zero coefficients of the extra variables occurring in ψ'_i to be bounded away from 0 and ∞ , independently of N . These two points are critical to the application in section 4: the first allows averaging of the k extra variables over a convex region contained in $[-O(N), O(N)]^k$, the second allows rescaling of the forms before the Cauchy-Schwarz applications without introducing extra factors depending on N . These issues do not arise in [10], the core reason being that the linear algebra algorithm we describe in Appendix B for constructing normal form extensions may be done over \mathbb{Q} , manipulating only with rationals of bounded height.

Before stating a proposition on the existence of such normal forms, let us define a suitable notion of non-degeneracy. This is very closely linked to the notion of finite Cauchy-Schwarz complexity, which one may find discussed in [10] and [8], but with a slightly technical modification to make the quantitative details easier to handle. Let $(\psi_1, \dots, \psi_m) = \Psi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and fix $i \in [m]$. We say that a partition \mathcal{P}_i of $[m] \setminus \{i\}$, i.e.

$$[m] \setminus \{i\} = \bigcup_{k=1}^{s+1} \mathcal{C}_k$$

for some disjoint sets \mathcal{C}_k and some $s \geq 0$, is ‘suitable’ if

$$\psi_i \notin \text{span}(\psi_j : j \in \mathcal{C}_k)$$

for any k . Then we define the ‘ \mathcal{P}_i -degeneracy variety’ $V_{\mathcal{P}_i}$ to be the set of all systems Ψ for which \mathcal{P}_i is not suitable. It is easy to observe that Ψ lies in the intersection of all the degeneracy varieties over all i if and only if some ψ_i is a real multiple of another ψ_j . We call this intersection ‘the degeneracy variety’ V_{degen} .

As we did for describing degeneracy properties of L , we let

$$\text{dist}(\Psi, V_{\mathcal{P}_i}) := \inf_{\Psi' \in V_{\mathcal{P}_i}} \|\Psi - \Psi'\|_\infty,$$

where the L^∞ norm is taken on \mathbb{R}^{mn} .

Proposition 11. *Let $(\psi_1, \dots, \psi_m) = \Psi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a system of real linear forms with coefficients bounded above in absolute value by C . Furthermore, suppose that there exists an absolute $c > 0$ such that for each i there exists some partition \mathcal{P}_i for which $\text{dist}(\Psi, V_{\mathcal{P}_i}) \geq c$. Then for each i there is an extension $\Psi' : \mathbb{R}^{n'} \rightarrow \mathbb{R}^m$ such that:*

- $n' \leq n + m - 1$.
- Ψ' is of the form

$$\Psi'(\mathbf{u}, w_1, \dots, w_{s+1}) := \Psi(\mathbf{u} + w_1 \mathbf{f}^1 + \dots + w_{s+1} \mathbf{f}^{s+1})$$

for some vectors $\mathbf{f}^k \in \mathbb{R}^n$ and some $s \geq 0$, such that $\|\mathbf{f}^k\|_\infty = O_{c,C}(1)$ for every k .

- Ψ' is in normal form with respect to ψ'_i .
- All the coefficients of ψ'_i attached to the variables w_k are $\Theta_{c,C}(1)$.

The proof is deferred to Appendix B. There we will also show why $\text{dist}(L, V_{\text{degen}}^*) \geq c$ implies that $\ker(L)$ admits a parametrisation Ψ such that for each i there exists a partition \mathcal{P}_i with $\text{dist}(\Psi, V_{\mathcal{P}_i}) = \Omega_{c,C}(1)$.

Let $s_i + 1$ be the minimal number of parts in a partition \mathcal{P}_i such that $\text{dist}(\Psi, V_{\mathcal{P}_i}) \geq c$, and $s = \max_i s_i$. We say that s is the c -Cauchy-Schwarz complexity of Ψ . It will follow from the arguments of section 4 that the von Neumann property (3) holds with $k = s + 1$.

We conclude this discussion of normal form by noting examples of bad behaviour in this real variable case. Indeed, take $\iota(N)$ some function which tends to infinity with N , and consider the forms

$$\psi_1(u_1, u_2) = (1 + \iota(N)^{-1})u_1 + u_2$$

$$\psi_2(u_1, u_2) = u_1 + u_2$$

Notice that Ψ tends to the degeneracy variety, so does not satisfy the hypotheses of the above proposition. One may of course run the algorithm from Lemma B.3 in Appendix B, and construct the normal form reparametrisation

$$\psi'_1(u_1, u_2, u_3, u_4) = (1 + \iota(N)^{-1})u_1 + u_2 + w_1$$

$$\psi'_2(u_1, u_2, u_3, u_4) = u_1 + u_2 + w_2.$$

Ψ does have all its non-zero coefficients bounded away from 0 and ∞ , but

$$\Psi'(u_1, u_2, w_1, w_2) = \Psi(u_1 + \iota(N)w_1 - \iota(N)w_2, u_2 - \iota(N)w_1 + (\iota(N) + 1)w_2),$$

so Ψ' is not obtained by bounded shifts of the u_i variables.

One final remark: a careful analysis of the proof in Appendix C of [10] demonstrates that it is sufficient for the general argument for Ψ merely to admit, for each i separately, an extension which is in normal form with respect to ψ_i . This is of little consequence in [10], as the simple algorithm which constructs normal form extensions with respect to a fixed ψ_i may easily be iterated. Yet certain quantitative aspects of such an iteration, critical to our analogous application of these ideas, are not immediately clear to us. We have stated Proposition B.3 for normal forms only with respect to a single i , in order to avoid this technical annoyance.

3. TRANSFERRING FROM \mathbb{Z} TO \mathbb{R}

In this section we prove Theorem 1, assuming the result of Theorem 25. This is done by splitting into two cases: a ‘quasi-rational’ case, in which L may be very well approximated by a matrix with rational coefficients of bounded height, and a ‘non-quasi-rational’ case, which deals with all other matrices L using Theorem 25. The first case may be understood using the classical rational linear forms theorem, whereas the second case is understood by transferring the problem from the setting of integer-variable functions into the setting of real-variable functions.

Let $f_1, \dots, f_d : [N] \rightarrow [-1, 1]$ be arbitrary functions, and assume that

$$|T_\varepsilon(f_1, \dots, f_d)| \geq \delta.$$

From the previous section we know that L possesses a rank matrix M , which for notational convenience we suppose consists of the first m columns of L . Then, we may factor $L = MS$, where S is an m -by- d matrix with coefficients that are $O_{c,C}(1)$, and whose first m -columns form the identity matrix. For j in the range $m+1 \leq j \leq d$, let

$$\mathbf{v}_j := M^{-1} \begin{pmatrix} \lambda_{1j} \\ \vdots \\ \lambda_{mj} \end{pmatrix}$$

denote the vector $\mathbf{v}_j \in \mathbb{R}^m$ which forms the j^{th} column of S .

The treatment of the non-quasi-rational case will require recourse to standard notions of equidistribution, applied to the sequences $(\mathbf{v}_j n_j)_{n_j=1}^N$. The factorisation $L = MS$ is not strictly necessary for the progress of the proof, but it will clarify matters considerably to restrict ourselves to the familiar world of equidistribution with respect to the quotient space $\mathbb{R}^m / \mathbb{Z}^m$, rather than \mathbb{R}^m quotiented by some other more mysterious lattice.

Before embarking upon the proof of Theorem 1 in these two cases, we note that we haven’t yet proved that the quantity δ is bounded in terms of N . We do that now, as this fact will be a useful background result in the more intricate estimations to follow.

Proposition 12. *We have*

$$T_\varepsilon(1, \dots, 1) \ll_{c,C,\varepsilon} 1.$$

Proof. By using the factorisation $L = MS$, one sees that $N^{d-m} T_\varepsilon(1, \dots, 1)$ is the number of solutions $n_1, \dots, n_d \leq N$ to

$$\begin{pmatrix} n_1 \\ \vdots \\ n_m \end{pmatrix} + \sum_{j=m+1}^d \mathbf{v}_j n_j \in M^{-1} B_\varepsilon^m.$$

Fixing a choice of the variables n_{m+1}, \dots, n_d forces the vector $(n_1, \dots, n_m)^T$ to lie in a convex region of diameter $O_{c,C,\varepsilon}(1)$. There are at most $O_{c,C,\varepsilon}(1)$ such points, and so in total we have the claimed bound. \square

We now recall various aspects of the qualitative theory of equidistribution.

Definition 13. Let $(x_n)_{n=1}^\infty$ be a sequence in some compact metric space X . We say that $(x_n)_{n=1}^\infty$ is asymptotically equidistributed if

$$\frac{1}{N} \sum_{n=1}^N f(x_n) \rightarrow \int_X f(x) d\mu(x)$$

as $N \rightarrow \infty$, for all continuous functions $f : X \rightarrow \mathbb{C}$, where μ is a fixed Borel probability measure.

The issue of asymptotic equidistribution is well understood by the multidimensional Weyl's criterion ([21, Proposition 1.1.2.], say). From this the next result follows easily – consider [21, Proposition 1.1.5.] and the subsequent remark, for example.

Theorem 14. If $\theta \in \mathbb{R}^d$ and $\mathbf{x}_n = n\theta \bmod \mathbb{Z}^d$ is a linear sequence then x_n asymptotically equidistributes in a finite union of cosets of a particular subtorus $T \leq \mathbb{T}^d$. Furthermore, if θ has some irrational coordinate then the subtorus T is non-trivial.

The critical observation in the non-quasi-rational case will be that, if B is a convex region, the number of solutions to $L\mathbf{n} \in B$ is not overly-sensitive to small perturbations to B . If one were to consider the simpler setting of a fixed matrix L , with N tending to infinity, Theorem 14 would be a perfectly adequate tool. The only non-quasi-rational requirement would be for S to have at least one irrational entry. To avoid repetition we will not describe this argument in detail, as it is a strict simplification of the arguments to come in this section, yet, if the reader finds the forthcoming quantitative technicalities distasteful, we encourage them to first consider this simpler scenario.

When one considers a large (but fixed) scale N , if S is within $O(N^{-1})$ of a rational matrix with bounded coefficients then the sequences $(\mathbf{v}_j n)_{n=1}^N$ do not have time to equidistribute. As coefficients of L may depend on N , we will need to make use of the following Theorem 16, which gives a quantitative decomposition of a linear phase into rational and equidistributing parts (with a slowly varying error term). This theorem may be found as [5, Theorem 6].

Before stating this result, we will need to recall certain definitions from [5] and [12], concerning rational approximations. These definitions are not completely standard, but are extremely useful.

Definition 15. For a natural number N and a positive quantity A , we say that $\theta \in \mathbb{R}^m$ is:

- (A, N) -smooth if $\|\theta\|_{\mathbb{T}^m} \leq \frac{A}{N}$,
- A -rational if $q\theta \in \mathbb{Z}^m$ for some $q \leq A$,
- (A, N) -irrational if $\|\mathbf{q} \cdot \theta\|_{\mathbb{T}} \geq \frac{A}{N}$ for all $\mathbf{q} \in \mathbb{Z}^m \setminus \{0\}$ with $\|\mathbf{q}\|_1 \leq A$.

We also need to consider sub-tori of \mathbb{T}^m in a quantitative sense. Indeed, given a subtorus $T \leq \mathbb{T}^m$ there is a matrix in $J \in \text{SL}_m(\mathbb{Z})$ such that $J(T) = \mathbb{T}^{m'} \times \{0\}^{m-m'}$. We say that T has complexity at most A if there is such a matrix J whose coefficients have absolute values at most A .

Finally, we say that θ is (A, N) -irrational in T if $\theta \bmod \mathbb{Z}^m \in T$, and $J(\theta)$ is (A, N) -irrational in $\mathbb{T}^{m'}$ in the sense of Definition 15.

We now state the key theorem.

Theorem 16. [5, Theorem 6] Given $\theta \in \mathbb{R}^m$, a positive integer N , a constant A_0 , and a growth function \mathcal{F} . there is a quantity A satisfying $A_0 \leq A \ll_{m, \mathcal{F}, A_0} 1$ and a decomposition

$$\theta = \theta_{\text{smth}} + \theta_{\text{rat}} + \theta_{\text{irrat}}$$

such that

- (1) $\|\boldsymbol{\theta}_{\text{smth}}\|_\infty \leq \frac{A}{N}$. In particular, $\boldsymbol{\theta}_{\text{smth}}$ is (A, N) -smooth;
- (2) $\boldsymbol{\theta}_{\text{rat}}$ is A -rational;
- (3) $\boldsymbol{\theta}_{\text{irrat}}$ is $(\mathcal{F}(A), N)$ -irrational in some subtorus $T \leq \mathbb{T}^m$ of complexity at most A .

The proof of Theorem 16 is a relatively straightforward iteration procedure, decreasing the dimension of T at each step. The ability to choose an arbitrary growth function \mathcal{F} , a by-now-familiar upshot of regularity-type iterative procedures, is of course extremely helpful. In our application we will choose $\mathcal{F}(A) = A^{C_2}$, for some suitably large constant C_2 .

Let us now introduce the appropriate quantitative notions of equidistribution, to demonstrate this intended application.

Definition 17 (See [12]). *Given a length $N > 0$ and an error tolerance $\tau > 0$, a sequence $(\mathbf{x}_n)_{n=1}^N$ in \mathbb{T}^m is τ -equidistributed on \mathbb{T}^m if*

$$\left| \frac{1}{N} \sum_{n=1}^N F(\mathbf{x}_n) - \int_{\mathbb{T}^m} F(\mathbf{x}) d\mu(\mathbf{x}) \right| \leq \tau \|F\|_{\text{Lip}}$$

for all Lipschitz functions $F : \mathbb{T}^m \rightarrow \mathbb{C}$, where

$$\|F\|_{\text{Lip}} = \|F\|_\infty + \sup_{\substack{\mathbf{x} \neq \mathbf{y} \\ \mathbf{x}, \mathbf{y} \in \mathbb{T}^m}} \frac{|F(\mathbf{x}) - F(\mathbf{y})|}{\|\mathbf{x} - \mathbf{y}\|_{\mathbb{T}^m}}.$$

The issue of τ -equidistribution is considered in the following theorem from [12] (which is proved by a relatively standard application of the Fejér kernel).

Theorem 18. [12, Theorem 3.1] *Let $m \geq 1$, let $0 < \tau < \frac{1}{2}$, and let $\boldsymbol{\theta} \in \mathbb{R}^m$. Then, if the sequence $(\boldsymbol{\theta}n \pmod{\mathbb{Z}^m})_{n=1}^N$ is not τ -equidistributed in \mathbb{T}^m , there exists a $\mathbf{q} \in \mathbb{Z}^m$ with $0 < \|\mathbf{q}\|_1 < \tau^{-O_m(1)}$ such that $\|\mathbf{q} \cdot \boldsymbol{\theta}\|_{\mathbb{T}} \ll_m \frac{\tau^{-O_m(1)}}{N}$.*

Viewing this theorem in the contrapositive, the upshot is the following immediate corollary.

Corollary 19. *Let A, C_1 be constants, and suppose $\boldsymbol{\theta} \in \mathbb{R}^m$ is (A^{C_2}, N) -irrational, for some constant C_2 which is large enough (depending on C_1 and m). Then $(\boldsymbol{\theta}n)_{n=1}^N$ is A^{-C_1} -equidistributed in \mathbb{T}^m .*

This Corollary will be our major tool. In fact we will need to develop the theory for general subtori $T \leq \mathbb{T}^m$, but this mundane technical matter will be consigned to Appendix C.

Definition of quasi-rational

Motivated by the discussion above, we define what it means for the matrix L to be quasi-rational. Let C_1 be a large constant depending on ε , c and C , and set $A_0 = \max(C_1, C_1\delta^{-2})$, and let \mathcal{F} be a growth function of the form $\mathcal{F}(A) = A^{C_2}$ for some suitably large constant C_2 . For each \mathbf{v}_j we have the value A_j satisfying $A_0 \leq A_j \leq_{\delta, \mathcal{F}} 1$ and the decomposition

$$\mathbf{v}_j = \mathbf{v}_{j, \text{smth}} + \mathbf{v}_{j, \text{rat}} + \mathbf{v}_{j, \text{irrat}}$$

provided by Theorem 16. We will let $A := \max(A_j : m+1 \leq j \leq d)$. If all the $\mathbf{v}_{j, \text{irrat}}$ are zero, we say that L is quasi-rational. Otherwise, we say that L is non-quasi-rational.

3.1. Quasi-rational case. Our proof for this case will involve reducing to a result which although immediate from the methods of [10] is not stated in that paper (the focus being on results over primes). For ease of reference we will state the required theorem, and indicate how one may construct a proof from an appropriate subset of the work of [10].

Theorem 20. *Let L be an m -by- d matrix with integer coefficients, $\|L\|_\infty \leq C$, and $\mathbf{b} \in \mathbb{R}^m$ some vector with $\|\mathbf{b}\|_\infty \leq C_1 N$. Let $K \subset [-N, N]^d$ be convex. Let $f_1, \dots, f_d : [N] \rightarrow \mathbb{C}$ be arbitrary functions with $\|f_i\|_\infty \leq 1$. Then*

$$\frac{1}{N^{d-m}} \sum_{\substack{\mathbf{n} \in \mathbb{Z}^d \cap K \\ L\mathbf{n} = \mathbf{b}}} \prod_{i=1}^d f_i(n_i) \ll_{C, C_1} \min \kappa(\rho) + o_\rho(1),$$

provided that, for some $k \leq d-1$, we have $\min_i \|f_i\|_{U^k[N]} \leq \rho$.

Proof Sketch. This theorem may be proved by following the proof of Theorem 1.8 of [10]. One follows the same linear algebraic reductions as those used in section 4 of [10] to reduce Theorem 1.8 to Theorem 7.1 of the same paper (the classical Generalised von Neumann Theorem). Theorem 7.1 may then be considered solely in the case of bounded functions (rather than in the more general case functions bounded by a pseudorandom measure).

This settles Theorem 20, with $k = s+1$, where s is the Cauchy-Schwarz complexity of the system of forms (ψ_1, \dots, ψ_d) parametrisng $\ker(L)$. But s is at most $d-2$, whence $k \leq d-1$ and Theorem 20 is proved. \square

We now reduce the quasi-rational case to this theorem. Let $S = S_{\text{rat}} + S_{\text{smth}}$ be the decomposition given by the application of Theorem 16 to all the columns \mathbf{v}_j . Without loss of generality we may assume that the first m columns of S_{rat} are the identity matrix. For simplicity of notation, we let D equal the convex set $M^{-1}(B_\varepsilon^m)$.

One could proceed by splitting $[N]^d$ into smaller boxes, over which the vector $S_{\text{smth}}\mathbf{n}$ varies little, but this is an unnecessary complication⁷. It is enough just to note that the image $S_{\text{smth}}([0, N]^d)$ is contained in an L^∞ -ball diameter $O(A)$.

Let Λ be the lattice $S_{\text{rat}}(\mathbb{Z})$. We see that $\Lambda \leq \frac{1}{q}\mathbb{Z}^d$ for some $q \leq A^{O(1)}$. Observe that $S\mathbf{n} \in D$ if and only if for some point $\boldsymbol{\lambda} \in \Lambda$ we have $S_{\text{rat}}\mathbf{n} = \boldsymbol{\lambda}$ and $S_{\text{smth}}\mathbf{n} \in D - \boldsymbol{\lambda}$. But there is only a solution to $S_{\text{smth}}\mathbf{n} \in D - \boldsymbol{\lambda}$ when $\boldsymbol{\lambda} \in D - S_{\text{smth}}([0, N]^d)$, which is a set of diameter $O_{c,C}(\varepsilon + A)$. There are at most $O_{c,C,\varepsilon}(A^{O(1)})$ points of $\frac{1}{q}\mathbb{Z}^d$ in $D - S_{\text{smth}}([0, N]^d)$, and hence at most $O_{c,C,\varepsilon}(A^{O(1)})$ points of Λ in $D - S_{\text{smth}}([0, N]^d)$. In other words, there exists a subset $\Lambda' \subset \Lambda$ of size $|\Lambda'| = O_{c,C,\varepsilon}(A^{O(1)})$ such that, for $\mathbf{n} \in [N]^d$, $S\mathbf{n} \in D$ only if $S_{\text{rat}}\mathbf{n} = \boldsymbol{\lambda}$.

In summary, we see

$$T_\varepsilon(f_1, \dots, f_d) = \sum_{\boldsymbol{\lambda} \in \Lambda'} \sum_{\substack{\mathbf{n} \in [N]^d \\ S_{\text{rat}}\mathbf{n} = \boldsymbol{\lambda} \\ S_{\text{smth}}\mathbf{n} \in D - \boldsymbol{\lambda}}} \prod_{i=1}^d f_i(n_i). \quad (5)$$

Consider just the inner sum of (5). The set of \mathbf{n} such that $\mathbf{n} \in [N]^d$ and $S_{\text{smth}}\mathbf{n} \in D - \boldsymbol{\lambda}$ is convex. Calling this domain K , which is contained in $[-N, N]^d$, the inner sum of (5) is

$$\sum_{\substack{\mathbf{n} \in K \\ S_{\text{rat}}\mathbf{n} = \boldsymbol{\gamma}}} \prod_{i=1}^d f_i(n_i) \quad (6)$$

⁷We only seek an upper bound, and all losses of $A^{O(1)}$ will be tame.

We wish to apply Theorem 20 to bound expression (6). By clearing denominators, we can assume that S_{rat} has integer coordinates and $\|S_{\text{rat}}\|_{\infty} \leq A^{O(1)}C$. Similarly we have $\|\lambda\|_{\infty} \ll_{c,C,\varepsilon} A^{O(1)}$. Applying Theorem 20, we conclude that (6) is

$$\ll_{c,C,\varepsilon} C(A)(\kappa(\rho) + o_{\rho}(1))$$

for some increasing function $C(A)$, provided that, for some $k \leq d-1$, we have $\min_i \|f_i\|_{U^k[N]} \leq \rho$.

Applying the triangle inequality to (5), we see that if $\min_i \|f_i\|_{U^k[N]} \leq \rho$ we have

$$C(A)^{-1}A^{-O(1)}\delta \ll_{c,C,\varepsilon} \kappa(\rho) + o_{\rho}(1). \quad (7)$$

Recall by construction that A satisfied $\delta^{-2} \ll A \ll_{\delta,\mathcal{F}} 1$, and so A is certainly some quantity which tends to infinity as δ tends to zero. Hence $C(A)^{-1}A^{-O(1)}\delta \geq F(\delta)$ for some function F which is $\kappa(\delta)$. Without loss of generality we may assume that F is strictly decreasing. The inverse function $F^{-1}(x)$ is also $\kappa(x)$, and recall that $\kappa(\kappa(x)) = \kappa(x)$. Therefore, applying F^{-1} to both sides of (7), we obtain

$$\delta \ll_{c,C,\varepsilon} \kappa(\rho) + o_{\rho}(1)$$

as required. So Theorem 1 is proved, provided L is quasi-rational.

We remark that in the analysis of (5) we have used heavily the fact that the error term $S_{\text{smth}}\mathbf{n}$ is a slowly varying *linear* function, rather than an arbitrary slowly varying function, in order to establish that the cut-off implied by the error function is convex.

3.2. Non-quasi-rational case. We set up as in the previous case. Let $S = S_{\text{smth}} + S_{\text{rat}} + S_{\text{irrat}}$ be the decomposition given by the application of Theorem 16 to all the \mathbf{v}_j . Without loss of generality we may assume that the first m columns of S_{rat} are the identity matrix, and for simplicity of notation, we once again let D equal the convex domain $M^{-1}(B_{\varepsilon}^m)$.

We now make a key definition. For each i , we extend f_i to a real variable function $\tilde{f}_i : \mathbb{R} \rightarrow [-1, 1]$ by defining

$$\tilde{f}_i(x) := \begin{cases} f_i(\langle x \rangle) & \text{if } \|x\|_{\mathbb{R}/\mathbb{Z}} < \eta \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

where η is some very small positive parameter, depending on δ via A , which will be chosen later.

The Gowers norms of the \tilde{f}_i over \mathbb{R} may easily be related to the Gowers norms of the f_i over N . Indeed, recall that

$$\|\tilde{f}_i\|_{U^k(\mathbb{R})}^{2^k} = \frac{1}{N^{k+1}} \int \prod_{(x,\mathbf{h}) \in \mathbb{R}^{k+1}} \prod_{\omega \in \{0,1\}^k} \tilde{f}_i(x + \mathbf{h} \cdot \omega) dx d\mathbf{h}.$$

The definition of \tilde{f}_i , combined with η being suitably small in terms of m , implies that the right-hand side is equal to

$$\frac{\text{vol}(R)}{N^{k+1}} \sum_{(x,\mathbf{h}) \in \mathbb{Z}^{k+1}} \prod_{\omega \in \{0,1\}^k} f_i(x + \mathbf{h} \cdot \omega), \quad (9)$$

where R is the region

$$R := \{(x', \mathbf{h}') \in \mathbb{R}^{k+1} : \forall \omega \in \{0,1\}^k, (x' + \mathbf{h}' \cdot \omega) \in [-\eta, \eta]\}.$$

We have $\text{vol}(R) \ll \eta^{2^k}$. Therefore, considering the definition of $\|f_i\|_{U^k[N]}$ in Appendix A, we have that expression (9) is $\ll \eta^{2^k} \|f_i\|_{U^k[N]}^{2^k}$.

To summarise, we have shown that

$$\|\tilde{f}_i\|_{U^k(\mathbb{R})} \ll \eta \|f_i\|_{U^k[N]}. \quad (10)$$

Example 21. We now present a sketch of the argument in the specific case of Example 3, i.e.

$$L = \begin{pmatrix} 1 & 0 & -\sqrt{2} & -1 + \sqrt{2} \\ 0 & 1 & -\sqrt{3} & -1 + \sqrt{3} \end{pmatrix}.$$

In this case, with $m = 2$ and $d = 4$, $S_{\text{smth}} = 0$ and S_{rat} is merely the identity matrix in the first m columns and zero elsewhere (i.e. $\mathbf{v}_{j,\text{smth}} = \mathbf{v}_{j,\text{rat}} = \mathbf{0}$ for all j where these vectors are defined, i.e. $m+1 \leq j \leq d$). Further, we have $S = L$. This example contains the main idea of the general case, but avoids the considerable technical complications.

The conclusion of Theorem 1 will follow immediately from the next claim.

Claim 22.

$$\frac{1}{N^2} \left| \sum_{\mathbf{n} \in B_\varepsilon^2} \prod_{i=1}^4 f_i(n_i) \right| \leq \frac{1}{(2\eta)^4 N^2} \left| \int_{\mathbf{x} \in B_\varepsilon^2} \prod_{i=1}^4 \tilde{f}_i(x_i) d\mu \right| + O(\varepsilon\eta) + \varepsilon\eta^{-1}o(1). \quad (11)$$

In other words, $\delta := T_\varepsilon(f_1, f_2, f_3, f_4) \leq \frac{1}{(2\eta)^4} \widetilde{T}_\varepsilon(f_1, f_2, f_3, f_4) + O(\varepsilon\eta) + \varepsilon\eta^{-1}o(1)$.

Sketch proof of Theorem 1 assuming Claim. Pick η to be a small multiple of δ , so that $O(\varepsilon\eta)$ term in (11) is at most $\frac{1}{2}\delta$. Using Claim 22, and the fact that δ is bounded, we conclude that

$$\delta^5 \ll_\varepsilon \widetilde{T}_\varepsilon(\tilde{f}_1, \tilde{f}_2, \tilde{f}_3, \tilde{f}_4) + o(1).$$

We may apply Theorem 7 to the \tilde{f}_i , deducing that

$$\delta^5 \ll_\varepsilon \kappa(\rho) + o_\rho(1), \quad (12)$$

provided that $\min_i \|f_i\|_{U^3(\mathbb{R})} \leq \rho$. But then, using the relation (10), (12) holds provided that $\min_i \|f_i\|_{U^k[N]} \leq \eta^{-1}\rho$. The parameter η is certainly less than 1, so (12) certainly holds provided $\min_i \|f_i\|_{U^k[N]} \leq \rho$. Taking the fifth-root of both sides, Theorem 1 is proved for this L . \square

It remains to prove Claim 22. In this proof, we recommend that the reader considers the manipulations as geometrically as possible, as it is from this viewpoint that the simple structure of the argument is most obvious.

Sketch proof of Claim. We have that the integral on the right-hand side of (11) is exactly equal to

$$\sum_{\mathbf{n} \in \mathbb{Z}^4} \left(\prod_{i=1}^4 f_i(n_i) \right) \mu([- \eta, \eta]^4 + \mathbf{n}) \cap L^{-1}B_\varepsilon^2), \quad (13)$$

where μ is Lebesgue measure. If

$$\|L\mathbf{n}\|_\infty \leq \varepsilon - C_1\eta,$$

for some large enough constant C_1 , then the set $[-\eta, \eta]^4 + \mathbf{n}$ is entirely contained within $L^{-1}B_\varepsilon^2$ and hence

$$\mu([- \eta, \eta]^4 + \mathbf{n}) \cap L^{-1}B_\varepsilon^2 = (2\eta)^4.$$

If

$$\|L\mathbf{n}\|_\infty \geq \varepsilon + C_1\eta,$$

then the set $[-\eta, \eta]^4 + \mathbf{n}$ is disjoint from $L^{-1}B_\varepsilon^2$ and hence

$$\mu([- \eta, \eta]^4 + \mathbf{n}) \cap L^{-1}B_\varepsilon^2 = 0.$$

For other n , the value of $\mu([- \eta, \eta]^4 + \mathbf{n}) \cap L^{-1}B_\varepsilon^2)$ lies somewhere between these two values. Therefore, using the fact that $\|f_i\|_\infty \leq 1$, we conclude that

$$\frac{1}{N^2} \sum_{\substack{\mathbf{n} \\ L\mathbf{n} \in B_\varepsilon^2}} \prod_{i=1}^4 f_i(n_i) = \frac{1}{(2\eta)^4 N^2} \int_{\substack{\mathbf{x} \\ L\mathbf{x} \in B_\varepsilon^2}} \prod_{i=1}^4 \tilde{f}_i(x_i) d\mu + \frac{1}{N^2} E, \quad (14)$$

where the error term E satisfies

$$|E| \ll \sum_{\substack{\mathbf{n} \in [N]^4 \\ \varepsilon - C_1 \eta \leq \|L\mathbf{n}\|_\infty \leq \varepsilon + C_1 \eta}} 1.$$

We now proceed to estimate this error term. Observe first that E may be considered as $|\{\mathbf{n} \in [N]^d : L\mathbf{n} \in R\}|$, for some compact domain R with volume depending on η . R is not convex, but we may cover R by small convex regions, say $O_{C_1}(\varepsilon\eta^{-1})$ convex sets R_i of the form $B_\eta^2 + \mathbf{a}$ for some shift \mathbf{a} .

Then the size of E may be upper-bounded by the number of solutions to

$$\begin{pmatrix} n_1 \\ n_2 \end{pmatrix} + \begin{pmatrix} -\sqrt{2} \\ -\sqrt{3} \end{pmatrix} n_3 + \begin{pmatrix} -1 + \sqrt{2} \\ -1 + \sqrt{3} \end{pmatrix} n_4 \in R_i \quad (15)$$

with $n_1, \dots, n_4 \leq N$, summed over each R_i . Provided η is small enough, the diameter of each R_i is at most $\frac{1}{10}$, and so, reducing modulo \mathbb{Z}^2 , we may certainly upper-bound E by the sum

$$\sum_{i \ll \varepsilon\eta^{-1}} \sum_{n_3, n_4 \leq N} 1_{R_i} \left(\begin{pmatrix} -\sqrt{2} \\ -\sqrt{3} \end{pmatrix} n_3 + \begin{pmatrix} -1 + \sqrt{2} \\ -1 + \sqrt{3} \end{pmatrix} n_4 \right), \quad (16)$$

where the regions R_i are now considered modulo \mathbb{Z}^2 .

The critical observation is that, due to the irrationality of the coefficients, the sum

$$\begin{pmatrix} -\sqrt{2} \\ -\sqrt{3} \end{pmatrix} n_3 + \begin{pmatrix} -1 + \sqrt{2} \\ -1 + \sqrt{3} \end{pmatrix} n_4$$

equidistributes over the torus $\mathbb{R}^2/\mathbb{Z}^2$. (This is immediate from Weyl's criterion, say). Therefore the expression (16) is at most

$$\sum_{i \ll \varepsilon\eta^{-1}} (\text{vol}(R_i) N^2 + o(N^2)).$$

Since $\text{vol}(R_i) \asymp \eta^2$, we have $|E| \ll \varepsilon\eta N^2 + \varepsilon\eta^{-1} o(N^2)$. Substituting this bound into equation 14 proves the claim. \square

We now proceed with the general case, in which there are three minor technical complications: the use of the rank matrix M to convert the question to one of equidistribution modulo \mathbb{Z}^m , the fact that equidistribution might take place not over \mathbb{T}^m but over a proper subtorus, and the requirement for the quantitative decomposition of S in the case where L depends on N . There is one more major complication, namely that if one of the rows of L is rational then the estimation of the error term E is not so straightforward: for certain unlucky values of ε , it might be that

$$\sum_{\substack{\mathbf{n} \in [N]^d \\ \varepsilon - C_1 \eta \leq \|L\mathbf{n}\|_\infty \leq \varepsilon + C_1 \eta}} 1 \gg T_\varepsilon(f_1, \dots, f_d)$$

for arbitrarily small η . We deal with all of these issues below, but the reader is encouraged to consider the rather long argument as merely a more intricate version of the previous proof.

Deduction of Theorem 1 in non-quasi-rational cases. Let \mathbf{v}_j be $(\mathcal{F}(A), N)$ -irrational in the subtorus $T_j \leq \mathbb{T}^m$, and let $T := T_{m+1} + \cdots + T_d$. Since L is non-quasi-rational, we may suppose that the dimension of T is at least 1. It is often useful to think instead of the dimension p of the annihilator of T , so we say that the dimension of T is equal to $m - p$ for some $0 \leq p \leq m - 1$. We let $\pi^{-1}(T) \subset \mathbb{R}^m$ denote the fibre lying above T , with respect to the natural projection π which projects modulo \mathbb{Z}^m . (In the case where $T = \mathbb{T}^m$, this fibre is of course all of \mathbb{R}^m , and various of the manipulations to come will be degenerate.)

Identifying the image of S

First, we identify more precisely the image set of S . Certainly $S_{\text{irrat}}\mathbf{n} \in \pi^{-1}(T)$, for all $\mathbf{n} \in \mathbb{Z}^d$. Then, since there exists $q \ll A^{O(1)}$ such that for all $\mathbf{n} \in \mathbb{Z}^d$ we have $S_{\text{rat}}(\mathbf{n}) \in \frac{1}{q}\mathbb{Z}^m$, we see that $(S_{\text{rat}} + S_{\text{irrat}})\mathbf{n}$ lies in the set $\pi^{-1}(T) + \Lambda$, for some finite $A^{-O(1)}$ -separated set Λ with size $|\Lambda| \leq A^{O(1)}$.

The set $\pi^{-1}(T)$ is the union of the cosets $T_0 + \mathbf{x}$ for $\mathbf{x} \in \mathbb{Z}^m$, of some linear subspace $T_0 \leq \mathbb{R}^m$ of dimension $m - p$. From Proposition C.1 we note that the complexity of T is at most $A^{O(1)}$, and this implies that $\text{dist}(T_0 + \mathbf{x}, T_0 + \mathbf{y}) \gg A^{-O(1)}$ for any two distinct cosets $T_0 + \mathbf{x}, T_0 + \mathbf{y}$. Therefore we may replace Λ by a new discrete set Λ' , albeit infinite, such that $\pi^{-1}(T) + \Lambda = T_0 + \Lambda'$. In other words, bringing these two paragraphs together, for all $\mathbf{n} \in [N]^d$ one has $(S_{\text{rat}} + S_{\text{irrat}})\mathbf{n} \in T_0 + \Lambda'$ and⁸ $\text{dist}(T_0 + \boldsymbol{\lambda}_1, T_0 + \boldsymbol{\lambda}_2) \geq A^{-O(1)}$ for all distinct pairs $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \Lambda'$.

We now replace Λ' by a finite subset Λ_0 . Indeed, similarly to the quasi-rational case, we may note that if $S\mathbf{n} \in D$ then $(S_{\text{rat}} + S_{\text{irrat}})\mathbf{n} \in (D - S_{\text{smth}}([0, N]^d)) \cap (T_0 + \Lambda')$. Since $\text{diam}(D - S_{\text{smth}}([0, N]^d)) = O_{c,C}(\varepsilon + A)$ there are only $O_{c,C,\varepsilon}(A^{O(1)})$ connected components of $T_0 + \Lambda'$ which intersect it. Therefore there exists $\Lambda_0 \subset \Lambda'$, a subset of at most $O_{c,C,\varepsilon}(A^{O(1)})$ points, such that, for $\mathbf{n} \in [N]^d$, $S\mathbf{n} \in D$ implies $(S_{\text{rat}} + S_{\text{irrat}})\mathbf{n} \in T_0 + \Lambda_0$.

We now consider the real pre-image $(S_{\text{rat}} + S_{\text{irrat}})^{-1}(T_0 + \Lambda')$: instead of upper-bounding $T_\varepsilon(f_1, \dots, f_d)$ by an expression involving $\widetilde{T}_\varepsilon(\widetilde{f}_1, \dots, \widetilde{f}_d)$, as in the above example, the upper bound will involve an integral of the functions \widetilde{f}_i over this pre-image. Noting that $S_{\text{rat}} + S_{\text{irrat}}$ is a matrix of full rank m , by considering the pre-image $(S_{\text{rat}} + S_{\text{irrat}})^{-1}(T_0 + \Lambda')$ we obtain a decomposition of \mathbb{Z}^d into

$$\mathbb{Z}^d = \bigcup_{\gamma \in \Gamma'} ((U_0 + \gamma) \cap \mathbb{Z}^d), \quad (17)$$

where Γ' is some discrete set and $U_0 \leq \mathbb{R}^d$ is the $d - p$ dimensional subspace $(S_{\text{rat}} + S_{\text{irrat}})^{-1}(T_0)$. The union is disjoint, and what's more (since $\|S_{\text{rat}} + S_{\text{irrat}}\|_{op} \ll_C 1$) we may ensure $\text{dist}(U_0 + \boldsymbol{\gamma}_1, U_0 + \boldsymbol{\gamma}_2) \gg_{c,C} A^{-O(1)}$ for all distinct pairs $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2 \in \Gamma'$.

By picking one pre-image per element of Λ_0 , one may similarly construct a subset $\Gamma_0 \subset \Gamma'$ of size $O_{c,C,\varepsilon}(A^{O(1)})$ such that

$$\{\mathbf{n} \in [N]^d : S\mathbf{n} \in D\} \subset \bigcup_{\gamma \in \Gamma_0} ((U_0 + \gamma) \cap \mathbb{Z}^d).$$

Comparison to an integral

We now come to use this information to relate $T_\varepsilon(f_1, \dots, f_d)$ to a real integral, on which (a technical modification of) Theorem 7 will be applicable.

The space $U_0 + \Gamma_0$ is a disjoint union of cosets of U_0 (to emphasise the geometry of the situation, we call each coset a *sheet*). The space $U_0 + \Gamma_0$ comes with a natural measure μ (either usual Lebesgue measure in the case where $U_0 + \Gamma_0 = \mathbb{R}^m$, or else with each

⁸If T has dimension m then $|\Lambda'| = 1$, so this statement is vacuously true.

sheet viewed as a isometric image of $\mathbb{R}^{d-p} \times \{0\}^p$ with Lebesgue measure of dimension $d-p$).

Now we can state the key claim, which is the general version of Claim 22.

Claim 23. *For some fixed constant $c_1 \asymp 1$, we have*

$$\frac{1}{N^{d-m}} \sum_{\mathbf{n} \in S^{-1}D} \prod_{i=1}^d f_i(n_i) = c_1 \frac{1}{\eta^{d-p}} \frac{1}{N^{d-m}} \int_{\substack{\mathbf{x} \in U_0 + \Gamma_0 \\ \mathbf{x} \in S^{-1}D}} \prod_{i=1}^d \tilde{f}_i(x_i) d\mu(\mathbf{x}) + O_{c,C,\varepsilon}(\eta A^{O(1)}). \quad (18)$$

Proof of Claim. The integral on the right hand side of equation (18) is exactly equal to

$$\sum_{\mathbf{n} \in U_0 + \Gamma_0} \left(\prod_{i=1}^d f_i(n_i) \right) \mu([-\eta, \eta]^d + \mathbf{n}) \cap S^{-1}D \cap (U_0 + \Gamma_0). \quad (19)$$

This is worth some explanation, as there is a minor technical subtlety. A priori, it is not obvious that a certain $\mathbf{x} \in U_0 + \Gamma_0$ which gives a non-zero contribution to the product $\prod_{i=1}^d \tilde{f}_i(x_i)$ in (18) couldn't have arisen from some $\mathbf{n} = \langle \mathbf{x} \rangle$ such that $\mathbf{n} \notin U_0 + \Gamma_0$. As previously argued, such an \mathbf{n} cannot satisfy $S\mathbf{n} \in D$, and so one may be concerned about the validity of the programme to relate the sum and the integral in expression (18). However, recall that in the case $p \geq 1$ (in the case $p = 0$ then $U_0 = \mathbb{R}^d$ and there is no issue) the different sheets of $U_0 + \Gamma'$ are distance $\Omega_{c,C}(A^{-O(1)})$ apart. In particular, by choosing η small enough, we can ensure that any integer $\mathbf{n} \notin U_0 + \Gamma_0$ satisfies $\text{dist}(\mathbf{n}, U_0 + \Gamma_0) > \eta$, and hence contributes nothing to the expression (19). Another useful upshot of this technical manoeuvre is that the set $([-\eta, \eta]^d + \mathbf{n}) \cap S^{-1}D \cap (U_0 + \Gamma_0)$ is restricted to a single sheet of $U_0 + \Gamma_0$.

We now analyse the μ term. The argument is analogous to that which followed expression (13), but here one must carefully consider the fact that dimension of U_0 may be less than d . When we use the boundary symbol $\partial(X)$ in the sequel, we are referring to the topological boundary of the space X when X is considered as a subset of a $d-p$ dimensional space.

Let $\mathbf{n} \in [N]^d \cap (U_0 + \Gamma_0)$. If both $\mathbf{n} \in S^{-1}D$ and $\text{dist}(\mathbf{n}, \partial(S^{-1}D \cap (U_0 + \Gamma_0))) > \eta$, the inclusion

$$([-\eta, \eta]^d + \mathbf{n}) \cap (U_0 + \Gamma_0) \subset (S^{-1}D \cap (U_0 + \Gamma_0))$$

holds. Hence

$$\mu([-\eta, \eta]^d + \mathbf{n}) \cap S^{-1}D \cap (U_0 + \Gamma_0) = \mu([-\eta, \eta]^d + \mathbf{n}) \cap (U_0 + \Gamma_0),$$

and by translation invariance this is equal to $\mu([-\eta, \eta]^d \cap (U_0 + \Gamma_0))$. This is a constant of size $\frac{1}{c_1} \eta^{d-p}$, for some $c_1 \asymp 1$. [Note once more that it is only U_0 itself which has a non-empty intersection with $[-\eta, \eta]^d$, and no other coset $U_0 + \gamma$.]

If $\mathbf{n} \notin S^{-1}D$ and $\text{dist}(\mathbf{n}, \partial(S^{-1}D \cap (U_0 + \Gamma_0))) > \eta$, then

$$\mu([-\eta, \eta]^d + \mathbf{n}) \cap S^{-1}D \cap (U_0 + \Gamma_0) = 0.$$

If $\text{dist}(\mathbf{n}, \partial(S^{-1}D \cap (U_0 + \Gamma_0))) \leq \eta$, the value $\mu([-\eta, \eta]^d + \mathbf{n}) \cap S^{-1}D \cap (U_0 + \Gamma_0)$ will take some value between the maximum $\frac{1}{c_1} \eta^{d-p}$ and 0.

Hence, using the fact that $\|f_i\|_\infty \leq 1$, we have that (19) is equal to

$$\frac{1}{c_1} \eta^{d-p} \sum_{\mathbf{n} \in S^{-1}D} \left(\prod_{i=1}^d f_i(n_i) \right) + \eta^{d-p} E, \quad (20)$$

where the error term E has size at most

$$\ll \sum_{\substack{\mathbf{n} \in [N]^d \cap (U_0 + \Gamma_0) \\ \text{dist}(\mathbf{n}, \partial(S^{-1}D \cap (U_0 + \Gamma_0))) \leq \eta}} 1. \quad (21)$$

Estimating the error in the comparison

Therefore, we see that Claim 23 will follow immediately from a suitable bound on E , which we give in the next proposition.

Proposition 24. *For certain suitable choices of η and the function \mathcal{F} , the error E is at most $O_{c,C,\varepsilon}(\eta A^{O(1)} N^{d-m})$*

Proof of Proposition. The proposition follows relatively straightforwardly from the equidistribution properties of S_{irrat} , although (as is the theme of the paper) the details are moderately technical.

Let R_η denote the image under the mapping of $S_{\text{rat}} + S_{\text{irrat}}$ of the domain of summation in (21), i.e. the image of $S_{\text{rat}} + S_{\text{irrat}}$ when applied to

$$Z := \{\mathbf{x} \in [0, N]^d \cap (U_0 + \Gamma_0) : \text{dist}(\mathbf{x}, \partial(S^{-1}D \cap (U_0 + \Gamma_0))) \leq \eta\}.$$

Identifying this set will help us better understand the set of points which E is counting. When $p = 0$ (and so $U_0 = \mathbb{R}^d$), one may think of R_η as an $O_{c,C}(\eta)$ -thickening⁹ of the boundary of D , as in our sketched example above. In the case of general p , however, it is not adequate to upper bound the set R_η this way¹⁰.

Covering the error with convex regions

We proceed as follows. Fix some sheet $U_0 + \gamma$, for some $\gamma \in \Gamma_0$. The portion of Z lying on this sheet, i.e. the set $Z \cap (U_0 + \gamma)$, is an $O(\eta)$ -thickening of the boundary of the convex set $S^{-1}D \cap (U_0 + \gamma)$, considered as a subset of a $d - p$ -dimensional space. The convex set $S^{-1}D \cap (U_0 + \gamma)$ has diameter $O_{c,C}(\varepsilon)$, and the image $(S_{\text{rat}} + S_{\text{irrat}})(S^{-1}D \cap (U_0 + \gamma))$ is another convex set of diameter $O_{c,C}(\varepsilon)$. This set is a subset of some $T_0 + \lambda$, an $m - p$ -dimensional space.

By standard facts about the boundaries of closed convex sets, an $O_{c,C}(\eta)$ -thickening of $(S_{\text{rat}} + S_{\text{irrat}})(S^{-1}D \cap (U_0 + \gamma))$, viewed as a subset of an $m - p$ -dimensional space, has volume $O_{c,C,\varepsilon}(\eta)$. Furthermore, it is certainly true that, provided the implied constants are large enough, this $O_{c,C}(\eta)$ -thickening contains the image $(S_{\text{rat}} + S_{\text{irrat}})(Z \cap (U_0 + \gamma))$.

Recombining the sheets, we conclude that the image R_η has volume $O_{c,C,\varepsilon}(A^{O(1)}\eta)$, viewed as an $m - p$ -dimensional set. Finally, we observe that we may cover R_η by $O_{c,C,\varepsilon}(A^{O(1)}\eta^{-(m-p-1)})$ convex sets R_i of diameter $O_{c,C}(\eta)$.

Reducing to an equidistribution problem

We now use this analysis to upper bound E . Indeed, the size of E may be upper-bounded by the number of solutions $(S_{\text{rat}} + S_{\text{irrat}})\mathbf{n} \in R_i$, with $\mathbf{n} \in [N]^d$, summed over all the R_i . In other words, by the number of solutions to

$$\begin{pmatrix} n_1 \\ \vdots \\ n_m \end{pmatrix} + \sum_{j=m+1}^d (\mathbf{v}_{j,\text{rat}} + \mathbf{v}_{j,\text{irrat}})n_j \in R_i. \quad (22)$$

⁹By an η -thickening of some convex domain $B \in \mathbb{R}^n$, we mean the set B union the set of points with distance at most η from the boundary of B .

¹⁰Indeed, we have been so careful to consider $U_0 + \Gamma_0$ throughout because, if one of rows of L is rational, D may have a flat side which coincides with one of the sheets of $T_0 + \Lambda_0$. This will mean that, no matter how small we pick η , any $O_{c,C}(\eta)$ -thickening of the boundary of D will contain a large proportion of some sheet of $T_0 + \Lambda_0$, and upper bounding E this way will fail.

Provided η is small enough in terms of c and C , the diameter of each R_i is at most $\frac{1}{10}$, and so, reducing modulo \mathbb{Z}^m , we may certainly upper-bound E by the sum

$$\sum_i \sum_{\substack{n_{m+1}, \dots, n_d \\ n_j \leq N \forall j}} 1_{R_i} \left(\sum_{j=m+1}^d (\mathbf{v}_{j,\text{rat}} + \mathbf{v}_{j,\text{irrat}}) n_j \right), \quad (23)$$

where the regions R_i are now considered modulo \mathbb{Z}^m . Recall that the size of the i summation is $O_{c,C,\varepsilon}(A^{O(1)}\eta^{-(m-p-1)})$.

We perform the standard elimination of the $\mathbf{v}_{j,\text{rat}}$ terms. For some $q_j \leq A_j$ we have $q_j \mathbf{v}_{j,\text{rat}} = 0 \pmod{\mathbb{Z}^m}$. For each j , we split $[N]$ into arithmetic progressions modulo q_j . Let Ω denote the domain

$$\prod_{j=m+1}^d (\mathbb{N} \cap [0, q_j - 1]).$$

Then for each $\mathbf{r} \in \Omega$ there is a shift of R_i (which for simplicity we denote again by R_i) such that (23) is equal to

$$\sum_i \sum_{\mathbf{r} \in \Omega} \sum_{\substack{n_{m+1}, \dots, n_d \\ n_j \leq \frac{1}{q_j}(N-r_j) \forall j}} 1_{R_i} \left(\sum_{j=m+1}^d (n_j q_j \mathbf{v}_{j,\text{irrat}}) \right). \quad (24)$$

The influence of vector $\mathbf{v}_{j,\text{rat}}$ has now been removed.

We seek an upper bound, so we may certainly replace $\frac{N-r_j}{q_j}$ in the inner sum by $\frac{N}{q_j}$. Replacing R_i by another shift as necessary, we may upper-bound this sum further by

$$\sum_{\substack{n_{m+1}, \dots, n_d \\ n_j \leq N \forall j}} 1_{R_i} \left(\sum_{j=m+1}^d n_j q_j \mathbf{v}_{j,\text{irrat}} \right), \quad (25)$$

and it is to this sum that we will apply the quantitative theory of equidistribution.

Applying a Lipschitz approximation, and choosing parameters

We claim that the size of (25) is at most $O_{c,C}(N^{d-m}\eta^{m-p})$. Indeed, first observe that the sum is unchanged if we replace R_i with $R_i \cap T$. Letting μ_T be the natural probability measure¹¹ on the torus T , then $\mu_T(R_i \cap T) = O_{c,C}(\eta^{m-p})$. Further, provided that $\eta \ll_{c,C} A^{-O(1)}$, for any $0 < \sigma < \frac{1}{2}$ (to be chosen specifically later) we may take a decomposition

$$1_{R_i \cap T} = F_\sigma + O(G_\sigma)$$

given Lemma D.2, where

$$F_\sigma, G_\sigma : T \longrightarrow [0, 1]$$

are Lipschitz functions on T with Lipschitz constant at most $O(A^{O(1)}\frac{1}{\sigma})$, and such that

$$\int_{\mathbf{x} \in T} G_\sigma(\mathbf{x}) d\mu_T(\mathbf{x}) \ll \sigma.$$

We are now finally in a position to apply Proposition C.1 – the slight technical generalisation of Corollary 19 – to the phases $q_j \mathbf{v}_{j,\text{irrat}}$. Pick growth function $\mathcal{F}(A) = A^{C_2}$, where C_2 is some constant which is much larger than any of the $O_m(1)$ constants which have occurred already (in our analysis of the gaps between different sheets of T_0 , in the application of Lemma D.2, etc.). Then, by the above decomposition and

¹¹See Appendix C for more detail.

Proposition C.1, we immediately establish that (25) is at most a constant (depending on c , C , *but not on* \mathcal{F}) times

$$N^{d-m} \left(\int_{\mathbf{x} \in T} 1_{R_i \cap T} d\mu_T(\mathbf{x}) + \frac{A^{-C_3}}{\sigma} + \sigma \right),$$

for a large constant C_3 . We choose $\sigma = A^{-\frac{1}{2}C_3}$ to balance the error terms.

Finally, recall that $\mu_T(R_i \cap T) = O_{c,C}(\eta^{m-p})$. Therefore, let us pick

$$\eta = A^{-\frac{1}{2(m-p)}C_3}.$$

Then η is smaller than all the previous $\Omega_{c,C}(A^{-O_m(1)})$ terms, as we required earlier in the proof, but we have also established that (25) is at most $O_{c,C,\varepsilon}(N^{d-m}\eta^{m-p})$ as claimed.

Summing over all the $\mathbf{r} \in \Omega$ and all the regions R_i , we conclude that the size of E is $O_{c,C,\varepsilon}(\eta A^{O(1)} N^{d-m})$ as claimed. \square

We have successfully bounded the error E , and, as remarked before, this resolves Claim 23. \square

Finishing the derivation

To finish the derivation of Theorem 1, we perform manipulations with (18) which are very similar to those from both the example and the quasi-rational case. Indeed, since we guaranteed that $A \geq C_1 \delta^{-2}$ by construction, the error is certainly at most $\frac{1}{2}\delta$, and hence

$$\frac{1}{N^{d-m}} \int_{\substack{\mathbf{x} \in U_0 + \Gamma_0 \\ \mathbf{x} \in S^{-1}D}} \prod_{i=1}^d \tilde{f}_i(x_i) d\mu(\mathbf{x}) \gg_{c,C,\varepsilon} \delta \eta^{d-p}.$$

We may then apply Theorem 25, a refinement of Theorem 7 to precisely this setting, to conclude that for all i there exists some $k \leq m+1$ such that

$$\kappa(\rho) + o_\rho(1) \gg_{c,C,\varepsilon} C(A) \delta \eta^{d-p}, \quad (26)$$

for some function $C(A)$, provided there is some f_i with $\|\tilde{f}_i\|_{U^k(\mathbb{R})} \leq \rho$.

Finally, using the relationship (10) between $\|\tilde{f}_i\|_{U^k(\mathbb{R})}$ and $\|f_i\|_{U^k[N]}$ established earlier, we conclude that

$$\kappa(\rho) + o_\rho(1) \gg_{c,C,\varepsilon} C(A) \delta \eta^{d-p}, \quad (27)$$

provided there is some f_i with $\|f_i\|_{U^k[N]} \leq \eta^{-1}\rho$. The parameter η is certainly less than 1, so (27) certainly holds provided there is some f_i with $\|f_i\|_{U^k[N]} \leq \rho$. The expression $\delta \eta^{d-p} C(A)$ is greater than some function $F(\delta)$ which is $\kappa(\delta)$ and strictly monotonically decreasing. Applying F^{-1} to both sides of (27) we get

$$\kappa(\rho) + o_\rho(1) \gg_{c,C,\varepsilon} \delta$$

provided there is some f_i with $\|f_i\|_{U^k[N]} \leq \rho$.

So, assuming Theorem 25, we have proved Theorem 1 in the non-quasi-rational case. \square

We remark that when all the f_i are non-negative, one may introduce a simpler upper-bounding procedure to convert from a discrete to a continuous solution count. Indeed, provided that η is small enough in terms of ε , we have

$$T_\varepsilon(f_1, \dots, f_d) \leq \frac{1}{\eta^d} \tilde{T}_{2\varepsilon}(\tilde{f}_1, \dots, \tilde{f}_d). \quad (28)$$

The expansion from ε to 2ε is of no consequence in the proof of Theorem 7, and so Theorem 1 follows immediately.

4. PROOF OF THE REAL VARIABLE VON NEUMANN THEOREM

In this section we prove Theorem 7, and therefore complete the proof of Theorem 1. We also prove a technical refinement of this theorem, which was required in section 3 when stating and proving Claim 23. The key requirement, we recall, was a bound on

$$\frac{1}{N^{d-m}} \int_{\substack{\mathbf{x} \in U_0 + \Gamma_0 \\ \mathbf{x} \in S^{-1}D}} \prod_{i=1}^d \tilde{f}_i(x_i) d\mu.$$

We make this a theorem:

Theorem 25. *Let D be some convex compact domain with diameter $O_{c,C}(\varepsilon)$, containing the origin, and let L satisfy the hypotheses of Theorem 7. Let $L = MS$ be the usual factorisation, and let $g_1, \dots, g_d : [0, N] \rightarrow [-1, 1]$ be arbitrary functions. Let subspace U_0 , discrete set Γ_0 , and parameter A be as defined in section 3. Then there is a $k \leq d-1$ such that, if $\min_i \|g_i\|_{U^k(\mathbb{R})} \leq \rho$, one has*

$$\frac{1}{N^{d-m}} \int_{\substack{\mathbf{x} \in U_0 + \Gamma_0 \\ \mathbf{x} \in S^{-1}D}} \prod_{i=1}^d g(x_i) d\mu \ll_{c,C,\varepsilon} \begin{cases} \kappa(\rho) + o_\rho(1) & \text{if } U_0 + \Gamma_0 = \mathbb{R}^d \\ C(A)(\kappa(\rho) + o_\rho(1)) & \text{otherwise, for some } C(A). \end{cases} \quad (29)$$

Observe that the first case of this theorem is exactly the statement of Theorem 7, and that the second case is exactly that which is used in section 3. Therefore a proof of Theorem 25 will settle both Theorem 1 and Theorem 7. The separation of the two cases is admittedly ugly, but when $U_0 + \Gamma_0 \neq \mathbb{R}^d$ one has the added complication that the kernel of S might be a linear complement to U_0 , and one needs to run the argument slightly differently.

Proof. We prove the second case of the Theorem, later mentioning the necessary simplifications to prove the first case. The outline of the proof is thus: we reduce the integral from the statement into an integral over a parametrisation of the kernel of $S_{\text{rat}} + S_{\text{irr}}$. Then we use Proposition B.3 to reparametrise in normal form, and finally apply the Cauchy-Schwarz inequality iteratively in order bound this integral by the Gowers norm.

Let us assume that

$$\frac{1}{N^{d-m}} \int_{\substack{\mathbf{x} \in U_0 + \Gamma_0 \\ \mathbf{x} \in S^{-1}D}} \prod_{i=1}^d g(x_i) d\mu = \delta.$$

By an averaging argument, and the fact that $|\Gamma_0| \ll_{c,C,\varepsilon} A^{O(1)}$, one may find some $\gamma \in \Gamma_0$ such that

$$\frac{1}{N^{d-m}} \int_{\substack{\mathbf{x} \in U_0 + \gamma \\ \mathbf{x} \in S^{-1}D}} \prod_{i=1}^d g(x_i) d\mu \gg_{c,C,\varepsilon} \delta A^{-O(1)}.$$

Without loss of generality, we assume $\|g_d\|_{U^k[N]} \leq \rho$. We also assume that N is sufficiently large in terms of c, C, ε and A . This will have the effect of introducing an $o(1)$ error term in the final reckoning.

Separating out the kernel

Let K be the kernel of $S_{\text{rat}} + S_{\text{irrat}}$. This is a vector space of dimension $d - m$, with an orthonormal basis $\{\mathbf{y}_1, \dots, \mathbf{y}_{d-m}\}$. We use this basis to define a parametrisation of the kernel

$$\begin{aligned} \Psi : \mathbb{R}^{d-m} &\longrightarrow K \\ \mathbf{x} &\mapsto (\psi_1(\mathbf{x}), \dots, \psi_d(\mathbf{x})) \end{aligned} \quad (30)$$

for certain real linear forms ψ_i . Concretely, letting $\mathbf{y}_{\mathbf{j}}^{(i)}$ denote the i^{th} coordinate of $\mathbf{y}_{\mathbf{j}}$ with respect to the standard basis, we have $\psi_i(\mathbf{x}) = \sum_{j=1}^{d-m} x_j \mathbf{y}_{\mathbf{j}}^{(i)}$. We have $\text{dist}(\Phi, V_{\text{degen}}) = \Omega_{c,C}(1)$, as is proved rigorously in Proposition B.5.

Now, extend this orthonormal basis $\{\mathbf{y}_1, \dots, \mathbf{y}_{d-m}\}$ for K to an orthonormal basis $\{\mathbf{y}_1, \dots, \mathbf{y}_{d-p}\}$ for U_0 , which we recall is a linear subspace of dimension $d - p$ for some $1 \leq p \leq m - 1$. (The $p = 0$ case is the first case of the Theorem, which we shall discuss later). Then, we may rewrite the assumption on δ as

$$\delta \ll_{c,C} A^{O(1)} \frac{1}{N^{d-m}} \int_{\mathbf{x}_1^{\mathbf{d}-\mathbf{p}}} 1_{D-S\gamma}(S(\sum_{j=1}^{d-p} x_j \mathbf{y}_{\mathbf{j}})) (\prod_{i=1}^d g'_i(\psi_i(\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}) + \sum_{j=d-m+1}^{d-p} x_j \mathbf{y}_{\mathbf{j}}^{(i)})) \quad (31)$$

where $\mathbf{x}_1^{\mathbf{d}-\mathbf{p}}$ is notation for integration over the entire variable collection x_1, \dots, x_{d-p} , and g'_i is a shift of g_i coming from the shift of U_0 by γ , i.e.

$$g'_i(u_i) := g_i(u_i + \gamma^{(i)}).$$

Observe that since we have chosen an orthonormal basis there is no contribution from a Jacobian, and that $|\gamma^{(i)}| = O_{c,C,\varepsilon}(A^{O(1)})$.

We wish to remove the presence of the variables $x_{d-m+1}, \dots, x_{d-p}$. Indeed, by construction we see that

$$1_{D-S\gamma}(S(\sum_{j=1}^{d-p} x_j \mathbf{y}_{\mathbf{j}})) = 1_{D-S\gamma}((S_{\text{rat}} + S_{\text{irrat}})(\sum_{j=d-m+1}^{d-p} x_j \mathbf{y}_{\mathbf{j}}) + S_{\text{smth}}(\sum_{j=1}^{d-p} x_j \mathbf{y}_{\mathbf{j}})).$$

Now, due to the restricted support of the g_i we may without loss of generality restrict the integral to variables $x_j \ll N$. But then, since $\|S_{\text{smth}}\|_{\infty} \leq \frac{A}{N}$ we have that $1_{D-S\gamma}(S(\sum_{j=1}^{d-p} x_j \mathbf{y}_{\mathbf{j}}))$ is zero unless $(S_{\text{rat}} + S_{\text{irrat}})(\sum_{j=d-m+1}^{d-p} x_j \mathbf{y}_{\mathbf{j}})$ lies in a certain convex region, containing the origin, of diameter $O_{c,C,\varepsilon}(A^{O(1)})$.

By applying Proposition B.6 to the matrix $S_{\text{rat}} + S_{\text{irrat}}$, this implies that $1_{D-S\gamma}(S(\sum_{j=1}^{d-p} x_j \mathbf{y}_{\mathbf{j}}))$ is zero unless the variables $x_{d-m+1}, \dots, x_{d-p}$ lie in some convex domain, containing the

origin, of diameter $O_{c,C,\varepsilon}(A^{O(1)})$.

We now apply an averaging argument to (31). Averaging over x_{d-m+1} to x_d , we conclude that there are some real shifts h_i of size $O_{c,C,\varepsilon}(A^{O(1)})$, and a vector $\mathbf{z} \in \mathbb{R}^m$ of size $\|\mathbf{z}\|_\infty = O_{c,C,\varepsilon}(A^{O(1)})$, such that

$$\delta \ll_{c,C,\varepsilon} A^{O(1)} \frac{1}{N^{d-m}} \int_{\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}} 1_D(\mathbf{z} + S_{\text{smth}}(\sum_{j=1}^{d-m} x_j \mathbf{y}_j)) \prod_{i=1}^d g'_i(\psi_i(\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}) + h_i). \quad (32)$$

Now let us remark how to proceed in the first case of the theorem, when $U_0 = \mathbb{R}^m$ (i.e. $p = 0$). In this case, we have $\Gamma_0 = \{0\}$, and in particular $|\Gamma_0| = 1$, so the first averaging argument does not appear. We then proceed by parametrising the kernel of S , rather than the kernel of $S_{\text{rat}} + S_{\text{smth}}$. The argument then runs identically until one considers $1_D(S(\sum_{j=1}^d x_j \mathbf{y}_j))$. In this instance

$$1_D(S(\sum_{j=1}^d x_j \mathbf{y}_j)) = 1_D(S(\sum_{j=d-m+1}^d x_j \mathbf{y}_j)),$$

and so one may pull this out of the inner integral completely, thus:

$$\delta \ll_{c,C} \int_{\mathbf{x}_1^{\mathbf{d}-\mathbf{m}+1}} 1_D(S(\sum_{i=d-m+1}^d x_i \mathbf{y}_i)) \frac{1}{N^{d-m}} \int_{\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}} (\prod_{i=1}^d g_i(\psi_i(\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}) + \sum_{j=d-m+1}^d u_j \mathbf{y}_j^{(i)})), \quad (33)$$

where $\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}$ is the variable collection x_1, \dots, x_{d-m} and $\mathbf{x}_{\mathbf{d}-\mathbf{m}+1}^{\mathbf{d}}$ is the variable collection x_{d-m+1}, \dots, x_d . We now average over x_{d-m+1}, \dots, x_d in a domain of diameter $O_{c,C}(\varepsilon)$ (applying Proposition B.6 as before). The upshot is that one once again reaches expression (32), but without the $A^{O(1)}$ factor and without the $1_D(\mathbf{z} + S_{\text{smth}}(\sum_{j=1}^{d-m} x_j \mathbf{y}_j))$ term. Further, the shifts h_i are now $O_{c,C}(\varepsilon)$.

Parametrising by normal form

Following this brief diversion, let us return to upper-bounding expression (32). Essentially all the constants from this point forward will depend on c, C, ε and A , so for readability we drop them from the notation. We now introduce the dummy variables which put the forms ψ_i into normal form. Indeed, let $R \subseteq [-O(N), O(N)]^{d-m}$ be the convex region such that $\mathbf{x}_1^{\mathbf{d}-\mathbf{m}} \in R$ if and only if both $\psi_i(\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}) + h_i$ is contained in the support of g'_i for all i and $\mathbf{z} + S_{\text{smth}}(\sum_{j=1}^{d-m} x_j \mathbf{y}_j) \in D$. R is contained in such a region since the vectors \mathbf{v}_j are orthonormal.

We apply Proposition B.3 to the forms ψ_i . Then, for *any* real numbers w_1, \dots, w_k , the right hand side of (32) is equal to

$$\frac{1}{N^{d-m}} \int_{\mathbf{x}_1^{\mathbf{d}-\mathbf{m}} + \sum_{j=1}^k w_j \mathbf{f}_j \in R} \prod_{i=1}^d g''_i(\psi'_i(\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}, w_1, \dots, w_k)) d\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}, \quad (34)$$

where $g''_i(x) := g'_i(x + h_i)$, the $\psi'_i : \mathbb{R}^{d-m+k} \rightarrow \mathbb{R}$ are certain linear forms with coefficients which are $O(1)$, the coefficients of the ψ'_d attached to the w_j variables are $\Omega(1)$, and $\mathbf{f}_1, \dots, \mathbf{f}_k$ are some vectors with $\|\mathbf{f}_j\|_\infty = O(1)$ for every j . Furthermore, the forms ψ'_i are in normal form with respect to ψ'_d , and $k \leq d-1$. Finally, an observation of the

proof of Proposition B.3 shows that we may without loss of generality create such an extension with $k \geq 2$.

Let

$$R' = \{(\mathbf{x}, w_1, \dots, w_k) \in \mathbb{R}^{d-m} \times [-N, N]^k : \mathbf{x} + \sum_{i=1}^k w_i \mathbf{f}_i \in R\}.$$

Since the vectors \mathbf{f}_i are bounded, we have

$$R' \subseteq [-O(N), O(N)]^{d-m+k}.$$

This is a critical point.

Writing \mathbf{w}_1^k for the variable collection w_1, \dots, w_k , we may write (34) as

$$\ll \frac{1}{N^{d-m+k}} \int_{\substack{\mathbf{x}_1^{d-m} \\ \mathbf{w}_1^k \in [-N, N]^k}} \prod_{i=1}^d g_i''(\psi'_i(\mathbf{x}_1^{d-m}, \mathbf{w}_1^k)) 1((\mathbf{x}_1^{d-m}, \mathbf{w}_1^k) \in R') d\mathbf{x}_1^{d-m} d\mathbf{w}_1^k. \quad (35)$$

Removing the convex cut-off

We now make some technical deductions, modifying the range of the w_i to be much greater than the range of the x_i , to enable the removing of x_i shifts at the end of the argument. We also replace R' by a Lipschitz cut-off, which then may be removed by a Fourier analysis argument.

The ψ'_i are in normal form with respect to ψ'_d . A further inspection of the proof of Proposition B.3 shows that, for each of the w_j variables, there is a non-empty collection of other forms $\{\psi'_i\}$ which are independent of the variable w_j . Furthermore we may partition the entire set $\{\psi'_i : i \neq d\}$ into such parts. Now, for each of the w_j variables we extend the range of integration to $[-C_1 N, C_1 N]$, where C_1 is some vast constant greater than all $O(1)$ constants appearing previously. This has no effect on the size (35), since we still have the $1((\mathbf{x}_1^{d-m}, \mathbf{w}_1^k) \in R')$ term, but soon this cut-off will be removed, and having a discrepancy in the relative sizes of certain variables will be very useful.

Next, Lemma D.1 tells us that for any small parameter σ we may write $1_{R'} = F_\sigma + O(G_\sigma)$, where F_σ and G_σ are non-negative Lipschitz functions supported on $[-O(N), O(N)]^{d-m+k}$ with Lipschitz constant $O(\frac{1}{\sigma N})$, both bounded pointwise by 1, and further with

$$\int_{\mathbf{x}} G_\sigma(\mathbf{x}) d\mathbf{x} = O(\sigma N^{d-m+k}).$$

Therefore we may replace the indicator function of R' by F_σ in (35), up to an error which is $O(\sigma)$. We will eventually choose σ to balance this error term with another.

Finally, with $e(t) := e^{2\pi i t}$ as usual, Lemma D.3 tells us that for any $X > 1$ we may write

$$F_\sigma(\mathbf{x}) = \int_{\|\boldsymbol{\xi}\|_\infty \leq X} c(\boldsymbol{\xi}) e(\boldsymbol{\xi} \cdot \mathbf{x}) d\boldsymbol{\xi} + O\left(\frac{1}{\sigma} \frac{\log X}{X}\right) \quad (36)$$

for some bounded function $c(\boldsymbol{\xi})$. The error this approximation contributes to (35) is $O\left(\frac{1}{\sigma} \frac{\log X}{X}\right)$. For our final choice of X (a slow growing function of ρ^{-1}) and for a suitable σ , this error will be also negligible. So, all told, we have that (up to an acceptable error) the expression (35) is at most a constant times

$$\frac{1}{N^{d-m+k}} \int_{\|\boldsymbol{\xi}\|_\infty \leq X} c(\boldsymbol{\xi}) \int_{\substack{\mathbf{x}_1^{d-m} \\ \mathbf{w}_1^k}}^* \prod_{i=1}^d g_i''(\psi'_i(\mathbf{x}_1^{d-m}, \mathbf{w}_1^k)) e(\boldsymbol{\xi} \cdot (\mathbf{x}_1^{d-m}, \mathbf{w}_1^k)) d\mathbf{u}_1^{d-m} d\mathbf{w}_1^k, \quad (37)$$

where the domain of the integration \int^* is

$$\begin{aligned}\mathbf{x}_1^{\mathbf{d}-\mathbf{m}} &\in [O(N), O(N)]^{d-m} \\ \mathbf{w}_1^{\mathbf{k}} &\in [-C_1N, C_1N]^k.\end{aligned}$$

Let us fix $\boldsymbol{\xi}$, and further let us fix the $\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}$ variables. In the setting of the finite group $\mathbb{Z}/N\mathbb{Z}$ the exponential phases $e(\boldsymbol{\xi} \cdot (\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}, \mathbf{w}_1^{\mathbf{k}}))$ can be absorbed simply by modifying the functions g_i'' by a suitable twist, the key point being that if $\boldsymbol{\xi} \cdot (\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}, \mathbf{w}_1^{\mathbf{k}})$ fails to be in the span of $\psi_i'(\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}, \mathbf{w}_1^{\mathbf{k}})$ then a Fourier expansion of g_i'' demonstrates that the expression corresponding to the inner integral of (37) is zero. This clean argument isn't quite so easy to apply here, but we may instead split up $e(\boldsymbol{\xi} \cdot (\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}, \mathbf{w}_1^{\mathbf{k}}))$ into a product of functions $\prod_{j=1}^k b_j$ such that a_j does not depend on the variable w_j . Since $k \geq 2$, this is clearly possible.

Applying Cauchy-Schwarz

Generalising the notation above, for $j \in [k]$ we let b_j denote any function which is bounded uniformly by 1 and which does not depend on the variable w_j . The exact value of b_j will vary from line to line, and further the exact variables which b_j does depend on (initially a subset of $(\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}, \mathbf{w}_1^{\mathbf{k}})$, but then in later iterations some additional shift variables) will be suppressed in the notation.

Then, the fact that the ψ_i' are in normal form with respect to ψ_d' means that we may partition the set $[d-1]$ into k parts \mathcal{C}_j , so that $i \in \mathcal{C}_j$ implies that $\psi_i'(0, \dots, 0, w_j, 0, \dots, 0) = 0$, i.e. ψ_i' is independent of the variable w_j . Therefore, combining this remark with the preceding two paragraphs, we are interested in upper bounding an expression of the form

$$\int_{\mathbf{w}_1^{\mathbf{k}} \in [-C_1N, C_1N]^k} g_d''(\psi_d'(\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}, \mathbf{w}_1^{\mathbf{k}})) \prod_{j=1}^k b_j d\mathbf{w}_1^{\mathbf{k}}, \quad (38)$$

with such an upper bound then integrated over the remaining variables: $\boldsymbol{\xi}$ and $\mathbf{x}_1^{\mathbf{d}-\mathbf{m}}$.

Observe by construction that

$$\psi_d'(0, \dots, 0, w_1, \dots, w_k) = w_1 + \dots + w_k.$$

[It would be enough that all the coefficients of w_j appearing in ψ_d' are $\theta(1)$, as one could then rescale.] Therefore (38) is at most

$$\ll \int_{\mathbf{w}_1^{\mathbf{k}} \in [-C_1N, C_1N]^k} g_d'''(w_1 + \dots + w_k) \prod_{j=1}^k b_j d\mathbf{w}_1^{\mathbf{k}}, \quad (39)$$

where g_d''' is a suitable $O(N)$ shift of g_d'' .

By applying Cauchy-Schwarz in each of the variables w_1 to w_k in turn, we establish that (39) is

$$\ll_{C_1} N^{k(1-\frac{1}{2^{k-1}})} \left(\int_{\substack{\mathbf{w}_1^{\mathbf{k}} \in [-C_1N, C_1N]^k \\ \mathbf{z}_1^{\mathbf{k}} \in [-C_1N, C_1N]^k}} \prod_{\boldsymbol{\alpha} \in \{0,1\}^l} g_d''' \left(\sum_{\substack{j \leq k \\ \alpha_j=0}} w_j + \sum_{\substack{j \leq k \\ \alpha_j=1}} z_j \right) d\mathbf{w}_1^{\mathbf{k}} d\mathbf{z}_1^{\mathbf{k}} \right)^{\frac{1}{2^k}}. \quad (40)$$

This expression may be immediately related to the real Gowers norm as given in Definition A.3. Indeed, let $h_j := z_j - w_j$ and $x := w_1 + \dots + w_k$. Performing this change of variables, we get (40) to be

$$\ll_{C_1} N^k \left(\frac{1}{N^{k+1}} \int \prod_{(\mathbf{x}, \mathbf{h}_1^k) \in D} \prod_{\alpha \in \{0,1\}^k} g_d'''(x + \alpha \cdot \mathbf{h}_1^k) dx d\mathbf{h}_1^k \right)^{\frac{1}{2^k}}, \quad (41)$$

where D is some domain which contains the box $[-C_1 N, C_1 N]^{k+1}$ (some new large C_1).

Since C_1 is large enough so that the range of integration of (x, \mathbf{h}_1^k) contains the entire domain on which the integrand is non-zero, one may extend the range of integration to all of \mathbb{R}^{k+1} . Transforming g_d''' to g_d by shifting, we conclude that (40) is at most

$$\ll N^k \|g_d\|_{U^k(\mathbb{R})},$$

noting that C_1 may now just be considered as just some large constant depending on C and c .

Finishing the proof

We recall the dependence of the constants on the various parameters, and include the $o(1)$ term which was suppressed by our initial assumption that N was large enough in terms of c, C, ε and A . Integrating (38) over ξ and all the $\mathbf{x}_1^{\mathbf{d}-\mathbf{k}}$ variables, and using our Lipschitz approximations, we have shown that (37) is at most

$$\ll_{c,C,\varepsilon,A} X^d \|g_d\|_{U^k(\mathbb{R})} + \frac{1}{\sigma} \frac{\log X}{X} + \sigma + o(1). \quad (42)$$

Therefore, recalling that $\|g_d\|_{U^k(\mathbb{R})} \leq \rho$ and picking σ to balance the error terms, we have

$$\delta \ll_{c,C,\varepsilon,A} X^d \rho + \left(\frac{\log X}{X} \right)^{\frac{1}{2}} + o(1). \quad (43)$$

Picking X to be a suitably slowly growing function of ρ^{-1} we get

$$\delta \ll_{c,C,\varepsilon,A} \kappa(\rho) + o(1)$$

This proves the second case of Theorem 25, indeed with the stronger conclusion that the $o(1)$ error term is independent of ρ . But recall that the manipulations for the first case, i.e. for the proof of Theorem 7, proceed almost identically but via (33), which does not have any dependence on the parameter A . Therefore, in the first case of Theorem 25, we have proved that

$$\delta \ll_{c,C,\varepsilon} \kappa(\rho) + o(1),$$

so Theorem 7 is also proved.

Combining this with the deduction in section 3, Theorem 1 is finally proved. \square

5. CONSTRUCTIONS

In this section we prove Theorem 6, the converse to Theorem 1, i.e. we show that requirement of being bounded away from the dual degeneracy variety V_{degen}^* is necessary for bounding $T_\varepsilon(f_1, \dots, f_n)$ by Gowers norms.

Indeed, suppose that $L = L(N)$ satisfies the hypotheses of Theorem 6. In particular

$$\liminf_{N \rightarrow \infty} \text{dist}(L, V_{\text{degen}}^*) = 0,$$

i.e. that $\text{dist}(L, V_{\text{degen}}^*) = \omega(N)^{-1}$, for some function $\omega(N)$ such that

$$\limsup_{N \rightarrow \infty} \omega(N) = \infty.$$

Let $\eta > 0$ be small¹², and pick N large so that $\omega(N) \geq \eta^{-1}$, and further such that $\eta N \geq \max(1, \varepsilon)$. We are going to construct functions f_1, \dots, f_d which, provided η is small enough, demonstrate that L fails to satisfy the von Neumann property. [All implied constants to follow will be independent of η .]

We begin by observing that $\|L\mathbf{n}\|_\infty \leq \varepsilon$ implies certain constraints on two of the variables n_i . Indeed, let $L' \in V_{\text{degen}}^*$ be the closest matrix in the dual degeneracy variety to L . By reordering columns, without loss of generality we may assume that there exist real numbers $\{a_i\}_{i=1}^m$ not all 0 s.t. for all $3 \leq j \leq d$ we have

$$\sum_{i=1}^m a_i \lambda'_{ij} = 0, \quad (44)$$

and further we may assume that $\lambda'_{i1} = \lambda_{i1}$ and $\lambda'_{i2} = \lambda_{i2}$ (else $L' \in V_{\text{degen}}^*$ is not the closest matrix to L). By reordering rows and rescaling, we may assume that a_1 has maximal absolute value amongst all the a_i , and that $|a_1| = 1$.

Define

$$b_1 := \sum_{i=1}^m a_i \lambda_{i1}, b_2 := \sum_{i=1}^m a_i \lambda_{i2},$$

and let $\mathbf{n} \in [N]^d$ be some solution to $\|L\mathbf{n}\|_\infty \leq \varepsilon$. The critical observation is that (44), combined with the assumptions on the a_i , implies that

$$|b_1 n_1 + b_2 n_2| \ll \eta N. \quad (45)$$

Indeed, for $3 \leq j \leq d$ we have

$$\left| \sum_{i=1}^m a_i \lambda_{ij} \right| = \left| \sum_{i=1}^m a_i (\lambda_{ij} - \lambda'_{ij}) \right| \ll \eta.$$

Since $\|L\mathbf{n}\|_\infty \leq \varepsilon$, we certainly have that

$$\left| b_1 n_1 + b_2 n_2 + \sum_{j=3}^d n_j \sum_{i=1}^m a_i \lambda_{ij} \right| \ll \varepsilon,$$

and then (45) follows by the triangle inequality and the fact that $\eta N \geq \varepsilon$.

The constraint (45) will turn out to be enough to show that L cannot satisfy the von Neumann property. We consider various cases, constructing different counterexample functions f_1 and f_2 based on the size and sign of b_1 and b_2 .

Case 1: b_1, b_2 both of the same sign, and $b_1, b_2 \gg_{c,C} 1$.

In this case, (45) implies¹³ that $n_1 \leq C_1 \eta N$ for some constant C_1 . Now, define $f_1 : [N] \rightarrow [-1, 1]$ to be the indicator function of the interval $[[C_1 \eta N], N] \cap \mathbb{N}$. We

¹²This parameter η has no relation to the η which was used in earlier sections.

¹³The same conclusion is true for n_2 , but this will not be needed.

then have

$$\begin{aligned} \|f_1 - 1\|_{U^k[N]} &\ll \left(\frac{1}{N^{k+1}} \sum_{x, h_1, \dots, h_k \ll C_1 \eta N} 1 \right)^{\frac{1}{2k}} \\ &\leq C_2 (C_1 \eta)^{\frac{k+1}{2k}} \end{aligned}$$

for some constant C_2 . However, observe that

$$\begin{aligned} |T_\varepsilon(f_1 - 1, 1, \dots, 1)| &= |T_\varepsilon(f_1, 1, \dots, 1) - T_\varepsilon(1, 1, \dots, 1)| \\ &= |0 - T_\varepsilon(1, 1, \dots, 1)| \gg_{c, C, \varepsilon} 1 \end{aligned}$$

by the hypotheses of Theorem 6. If L satisfied the von Neumann property, one would have

$$1 \ll_{c, C, \varepsilon} \kappa(C_2 (C_1 \eta)^{\frac{k+1}{2k}}) + o_\eta(1).$$

Picking η small enough, then N large enough, this inequality cannot possibly hold, and we have a contradiction. Therefore L does not satisfy the von Neumann property.

Case 2: b_1, b_2 of opposite signs, and $b_1, b_2 \gg_{c, C} 1$.

This is the most involved case, although the central idea is very simple. The condition (45) confines n_2 to lie within a certain length of a fixed multiple of n_1 . By constructing functions f_1 and f_2 using random choices of blocks of this length, but coupled in such a way that condition (45) is very likely to hold, we can guarantee that $T_\varepsilon(f_1 - p, f_2 - p, 1, \dots, 1)$ is bounded away from zero, where p is the probability chosen. However, despite the block construction and the coupling, the functions f_1 and f_2 still individually exhibit enough randomness to conclude that $\|f_1 - p\|_{U^k[N]} = o(1)$, and the same for f_2 , and so the von Neumann property cannot possibly hold.

We now fill in the technical details. Relation (45) implies that

$$|b_1 n_1 + b_2 n_2| \leq C_1 \eta N, \quad (46)$$

for some C_1 , and without loss of generality assume that b_1 is positive, b_2 is negative, and $|b_1|$ is at least $|b_2|$. Let C_2 be some parameter, chosen so that $(C_1 C_2 \eta)^{-1}$ is an integer. Such a C_2 will of course depend on η , but in magnitude we may pick $C_2 \asymp_{c, C} 1$. We consider the real interval $[0, N]$ modulo N , and for $x \in [0, N]$ and $0 \leq i \leq (C_1 C_2 \eta)^{-1} - 1$ we define the half-open interval modulo N

$$I_i := [x + i C_1 C_2 \eta N, x + (i + 1) C_1 C_2 \eta N).$$

This choice guarantees that

$$[0, N] = \bigcup_{i=0}^{(C_1 C_2 \eta)^{-1} - 1} I_i, \quad (47)$$

and the union is disjoint. Now, for δ a small constant to be chosen later¹⁴, we define

$$I_i^\delta := [x + (i + \frac{1}{2} - \delta) C_1 C_2 \eta N, x + (i + \frac{1}{2} + \delta) C_1 C_2 \eta N).$$

We will use the partition (47) to construct a function f_1 , using an averaging argument to choose an x so that the I_i^δ intervals capture a positive proportion of the solution density of the linear inequality system. Indeed, for $n_1 \in [N]$ let the weight $u(n_1)$ denote

¹⁴This δ is unrelated to the notation $\delta = T_\varepsilon(f_1, \dots, f_d)$ used in previous sections.

the number of $d - 1$ -tuples $n_2, \dots, n_d \leq N$ that together with n_1 satisfy the inequality $\|L\mathbf{n}\|_\infty < \varepsilon$. The weight $u(n_1)$ could be zero, of course. Then

$$\begin{aligned} \frac{1}{N} \int_0^N \sum_{n \in [N]} u(n) 1(n \in \cup_i I_i^\delta) dx &= \frac{1}{N} \sum_{n \in [N]} u(n) \int_0^N 1(n \in \cup_i I_i^\delta) dx \\ &= \sum_{n \in [N]} u(n) 2\delta \\ &= 2\delta N^{d-m} T_\varepsilon(1, \dots, 1) \end{aligned}$$

Therefore we may fix an x such that

$$\sum_{n \in [N]} u(n) 1(n \in \cup_i I_i^\delta) \gg_{c, C, \varepsilon} \delta N^{d-m} \quad (48)$$

Let us finally define the function f_1 . Let $p > 0$ be a small constant to be decided later. Then we define a random subset $A \subseteq [N]$ by picking all of $I_i \cap \mathbb{N}$ to be members of A , with probability p , or none of $I_i \cap \mathbb{N}$ to be members of A , with probability $1 - p$. We then make this same choice for each i satisfying $0 \leq i \leq (C_1 C_2 \eta)^{-1} - 1$ independently. Observe immediately that for each $n \in [N]$ the probability that $n \in A$ is always p (though these events are not always independent). We let $f_1(n)$ be the indicator function $1_A(n)$.

The function f_2 is defined in terms of f_1 . Indeed, let

$$J_i = \frac{b_1}{|b_2|} I_i \cap (0, N],$$

where the dilation is not considered modulo N but rather just as an operator on subsets of \mathbb{R} . Since $b_1 \geq |b_2|$ we have that these J_i also form a disjoint partition of $[0, N]$. [NB: If $b_1 > |b_2|$ it may be that certain J_i are empty, since the dilate of the corresponding I_i may land entirely outside $[0, N]$.] Then let B be the subset of $[N]$ defined so that for each i with J_i non-empty we have $J_i \cap \mathbb{N} \subseteq B$ if and only if $I_i \cap \mathbb{N} \subseteq A$. Note again that for each individual $n \in [N]$ the probability that $n \in B$ is always p . We let $f_2(n)$ be the indicator function $1_B(n)$.

Our first claim is that, in expectation, $T_\varepsilon(f_1, f_2, 1, \dots, 1)$ is significantly larger than $T_\varepsilon(p, p, 1, \dots, 1)$. Indeed, suppose that I_i is included in the set A , and suppose that $n_1 \in I_i^\delta$. Then, if $n_2 \in [N]$ satisfies $|\frac{b_1}{|b_2|} n_1 - n_2| \leq \frac{1}{|b_2|} C_1 \eta N$ and δ is small enough in terms of b_1 and b_2 , we certainly¹⁵ have $n_2 \in J_i$. Thus, by the observation (46), if $u(n_1)$ is at least 1, and if n_1 is contained in some I_i^δ interval included in A , then all n_2 which are part of solutions vectors featuring n_1 are contained in B . Therefore

$$\begin{aligned} \mathbb{E} T_\varepsilon(f_1, f_2, 1, \dots, 1) &= \frac{1}{N^{d-m}} \sum_{\substack{\mathbf{n} \in [N]^d \\ \|L\mathbf{n}\|_\infty \leq \varepsilon}} \mathbb{P}(n_1 \in A \wedge n_2 \in B) \\ &\geq \frac{1}{N^{d-m}} \sum_{\substack{\mathbf{n} \in [N]^d \\ \|L\mathbf{n}\|_\infty \leq \varepsilon}} \mathbb{P}(n_1 \in A \wedge n_1 \in I_i^\delta \text{ for some } i \wedge n_2 \in B) \\ &= \frac{1}{N^{d-m}} \sum_{n_1 \in [N]} u(n_1) p \mathbb{1}(n_1 \in I_i^\delta \text{ for some } i) \\ &\geq 2\delta p T_\varepsilon(1, \dots, 1) \end{aligned}$$

On the other hand $T_\varepsilon(p, p, 1, \dots, 1) = p^2 T_\varepsilon(1, \dots, 1)$, and hence

$$\mathbb{E} T_\varepsilon(f_1, f_2, 1, \dots, 1) - T_\varepsilon(p, p, 1, \dots, 1) \geq (2\delta p - p^2) T_\varepsilon(1, \dots, 1). \quad (49)$$

¹⁵This fact is the reason why we introduced the parameter δ .

Our second claim is that, in expectation, the Gowers norms $\|f_i - p\|_{U^k[N]}$ are small for $i = 1, 2$, $k \leq m + 1$. Consider first $i = 1$. Then

$$\mathbb{E}\|f_1 - p\|_{U^k[N]}^{2^k} \ll \frac{1}{N^{k+1}} \sum_{(x, \mathbf{h}) \in \mathbb{Z}^{k+1}} \mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k+1}} (f_1 - p\mathbb{1}_{[N]})(x + \mathbf{h} \cdot \omega) \right).$$

Observe that for fixed (x, \mathbf{h}) the random variables $(f_1 - p)(x + \mathbf{h} \cdot \omega)$ each have mean zero and, unless some two of the expressions $x + \mathbf{h} \cdot \omega$ lie in the same block I_i , these random variables are independent. Hence, apart from those exceptional cases, we may factor the expectation and conclude that

$$\mathbb{E} \left(\prod_{\omega \in \{0,1\}^{k+1}} (f_1 - p)(x + \mathbf{h} \cdot \omega) \right) = \prod_{\omega \in \{0,1\}^{k+1}} \mathbb{E}((f_1 - p)(x + \mathbf{h} \cdot \omega)) = 0.$$

Therefore, $\mathbb{E}\|f_1 - p\|_{U^k[N]}^{2^k}$ is

$$\begin{aligned} &\ll \frac{1}{N^{k+1}} \sum_{(x, \mathbf{h}) \in [-N, N]^{k+1}} \mathbb{1}(|\mathbf{h} \cdot (\omega_1 - \omega_2)| \leq C_1 C_2 \eta N \text{ for some } \omega_1 \neq \omega_2 \in \{0, 1\}^{k+1}) \\ &\ll \eta. \end{aligned}$$

Thus by Hölder's inequality we have $\mathbb{E}\|f_1 - p\|_{U^k[N]} \ll \eta^{\frac{1}{2^k}}$. The calculation for f_2 is essentially identical, noting that the length of the blocks J_i is also $O(\eta N)$.

It is possible that one could finish the argument here by considering a second-moment argument and choosing some explicit f_1 and f_2 . To avoid calculating a second moment, we argue as follows. Suppose for contradiction that L did satisfy the von Neumann property. Then

$$\begin{aligned} (2\delta p - p^2)T_\varepsilon(1, \dots, 1) &\leq |\mathbb{E}T_\varepsilon(f_1, f_2, 1, \dots, 1) - T_\varepsilon(p, p, 1, \dots, 1)| \\ &\ll |\mathbb{E}T_\varepsilon(f_1 - p, f_2, 1, \dots, 1)| + |\mathbb{E}T_\varepsilon(p, f_2 - p, 1, \dots, 1)| \\ &\ll \mathbb{E}(\kappa(\rho_1) + o_{\rho_1}(1)) + \mathbb{E}(\kappa(\rho_2) + o_{\rho_2}(1)), \end{aligned} \tag{50}$$

where ρ_1 (resp. ρ_2) is any chosen upper-bound on $\|f_1 - p\|_{U^k[N]}$ (resp. $\|f_2 - p\|_{U^k[N]}$).

We wish to reverse the order of the expectation operation and the κ and $o_{\rho_i}(1)$ functions, to conclude that the right-hand side is $\kappa(\eta) + o_\eta(1)$. We make two observations. Note first that

$$\mathbb{P}(\|f_1 - p\|_{U^k[N]} \gg \eta^{\frac{1}{2^{k+1}}}) \ll \eta^{\frac{1}{2^k}}$$

We choose the (random) upper-bound

$$\rho_1 \asymp \begin{cases} 1 & \text{if } \|f_1 - p\|_{U^k[N]} \gg \eta^{\frac{1}{2^{k+1}}} \\ \eta^{\frac{1}{2^{k+1}}} & \text{otherwise.} \end{cases}$$

Secondly, we may upper-bound the two $\kappa(\rho_i)$ functions by concave envelopes, so without loss of generality we may assume that the $\kappa(\rho_i)$ functions are concave.

Then by Jensen's inequality,

$$\begin{aligned} \mathbb{E}(\kappa(\rho_1) + o_{\rho_1}(1)) &\ll \kappa(\mathbb{E}\rho_1) + \mathbb{E}(o_{\rho_1}(1)) \\ &\ll \kappa(\eta^{\frac{1}{2^{k+1}}}) + o_\eta(1) \\ &\ll \kappa(\eta) + o_\eta(1). \end{aligned}$$

We do the same manipulation for f_2 . Picking p small enough, this yields a contradiction to (50) for η small enough and N large enough. Indeed, the left-hand side of (50) is $\Omega_{c,C,\varepsilon}(1)$.

Case 3: Exactly one of b_1, b_2 satisfies $b_i \gg_{c,C} 1$.

Without loss of generality we may assume that $b_1 \gg_{c,C} 1$. But then, as in Case 1, (45) implies that $n_1 \leq C_1 \eta N$ for some constant C_1 . The same construction as in Case 1 then applies.

Case 4: The remaining case. Under the assumptions of Theorem 6, no other case can in fact occur. To be precise, we assume in this case that both $|b_1|$ and $|b_2|$ are at most c_1 . But then consider the matrix L'' , defined by taking

$$\lambda''_{ij} = \lambda'_{ij}$$

when for all pairs $(i, j) \in [m] \times [d]$, except for $(1, 1)$ and $(1, 2)$. In these cases we let

$$\begin{aligned} \lambda''_{11} &= \lambda'_{11} - \frac{b_1}{a_1} \\ \lambda''_{12} &= \lambda'_{12} - \frac{b_2}{a_1}. \end{aligned}$$

Then

$$\sum_{i=1}^m a_i \lambda''_{ij} = 0$$

for all $1 \leq j \leq d$, i.e. $\|L - L''\|_\infty \leq \eta + c_1$ for some matrix L'' with rank less than m . Since $\eta + c_1 < c$, for c_1 small enough, this contradicts the assumptions of Theorem 6, so this final case does not arise.

We have covered all cases, and thus have concluded the proof of Theorem 6.

APPENDIX A. GOWERS NORMS

There are several existing expositions of the basic theory of Gowers norms – for example in [22], [7] and [10] – and the reader looking for an introduction to the theory in its full generality should certainly consult these references. However, in the interests of making the paper as self-contained as possible, we use this appendix to pick out the central definitions and notions which will be used in the main text.

Definition A.1. Let $N \in \mathbb{N}$. For a function $f : \mathbb{Z}/N\mathbb{Z} \rightarrow \mathbb{C}$, and an integer $d \geq 1$, define the Gowers U^d norm, $\|f\|_{U^d(N)}$ to be the unique non-negative solution to

$$\|f\|_{U^d(N)}^{2^d} = \frac{1}{N^{d+1}} \sum_{x, h_1, \dots, h_d} \prod_{\omega \in \{0,1\}^d} \mathcal{C}^{|\omega|} f(x + \mathbf{h} \cdot \omega), \quad (51)$$

where $|\omega| = \sum_i \omega_i$, and \mathcal{C} is the complex-conjugation operator.

For example,

$$\|f\|_{U^1(N)} = \frac{1}{N} \sum_x |f(x)|,$$

and

$$\|f\|_{U^2(N)} = \left(\frac{1}{N^3} \sum_{x, h_1, h_2} f(x) \overline{f(x + h_1)} \overline{f(x + h_2)} f(x + h_1 + h_2) \right)^{\frac{1}{4}}.$$

It is not immediately obvious that the right-hand side of (51) is always a non-negative real, nor why for $d \geq 2$ the U^d norms are genuine norms: proofs of both these facts may be found in [22].

The functions in the main text do not have a cyclic group as a domain but rather the interval $[N]$, but the theory may easily be adapted to this case.

Definition A.2. Let $N' \geq N$, $d \geq 1$, and $f : [N] \rightarrow \mathbb{C}$ any function. Identify $[N]$ with a subset of $\mathbb{Z}/N'\mathbb{Z}$ in the natural way. Then, we define the Gowers norm $\|f\|_{U^d[N]}$ to be the unique non-negative real solution to the equation

$$\|f\|_{U^d[N]}^{2^d} = \frac{1}{|R|} \sum_{x, h_1, \dots, h_d} \prod_{\omega \in \{0,1\}^d} \mathcal{C}^{|\omega|} f 1_{[N]}(x + \mathbf{h} \cdot \omega), \quad (52)$$

where $f 1_{[N]}$ is the extension by zero of f to $\mathbb{Z}/N'\mathbb{Z}$, the summation is over $x, h_1, \dots, h_d \in \mathbb{Z}/N'\mathbb{Z}$, and the set R is the set

$$R := \{x, h_1, \dots, h_d \in \mathbb{Z}/N'\mathbb{Z} : \text{for every } \omega \in \{0,1\}^d, x + \mathbf{h} \cdot \omega \in [N]\}.$$

It is worth remarking, though not directly needed to understand any statement in the main text, that one can immediately see that this definition is equivalent to

$$\|f\|_{U^d[N]} = \|f 1_{[N]}\|_{U^d(N')}/\|1_{[N]}\|_{U^d(N')}.$$

Taking $N' = O(N)$ we have $\|1_{[N]}\|_{U^d(N')} \asymp 1$, and thus $\|f\|_{U^d[N]} \asymp \|f 1_{[N]}\|_{U^d(N')}$, i.e. Definition A.1 and A.2 are comparable.

For us, the key note is that there is only a contribution to the summand in equation (52) when $|x| \leq N$ and for every i we have $|h_i| \leq N$. Further, it may be easily seen that $|R| \asymp N^{d+1}$. Therefore, if N'/N is sufficiently large we conclude that

$$\|f\|_{U^d[N]} \asymp \left(\frac{1}{N^{d+1}} \sum_{x, h_1, \dots, h_d} \prod_{\omega \in \{0,1\}^d} \mathcal{C}^{|\omega|} f(x + \mathbf{h} \cdot \omega) \right)^{\frac{1}{2^d}}. \quad (53)$$

This concludes our discussion of Gowers norms for functions on $[N]$.

In order to succinctly state Theorem 7 we referred to a Gowers norm $U^d(\mathbb{R})$. This is a less well-studied object, as the theory was developed (and has found its main applications) over finite groups. Nevertheless it may be perfectly well defined, and even deep aspects of its inverse theory may be deduced from the corresponding theory of the discrete Gowers norm. See [20].

Definition A.3. Let $f : \mathbb{R} \rightarrow \mathbb{C}$ be a bounded measurable function supported on $[0, 1]$, and let $d \geq 1$ be an integer. Then we define the Gowers norm $\|f\|_{U^d(\mathbb{R})}$ to be the unique non-negative real satisfying

$$\|f\|_{U^d(\mathbb{R})}^{2^d} = \int_{\mathbb{R}^{d+1}} \prod_{\omega \in \{0,1\}^d} \mathcal{C}^{|\omega|} f(x + \sum_{i=1}^d h_i \omega_i) dx dh_1 \cdots dh_d \quad (54)$$

where $|\omega| = \sum_i \omega_i$, and \mathcal{C} is the complex-conjugation operator.

For functions $g : [0, N] \rightarrow \mathbb{C}$, which are the type appearing in the main text, we define the function $f : [0, 1] \rightarrow \mathbb{C}$ by $f(x) := g(Nx)$, and then set

$$\|g\|_{U^d(\mathbb{R})} := \|f\|_{U^d(\mathbb{R})}.$$

Explicitly, a change of variables shows that

$$\|g\|_{U^d(\mathbb{R})}^{2^d} = \frac{1}{N^{d+1}} \int_{\mathbb{R}^{d+1}} \prod_{\omega \in \{0,1\}^d} \mathcal{C}^{|\omega|} g(x + \sum_{i=1}^d h_i \omega_i) dx dh_1 \cdots dh_d.$$

APPENDIX B. LINEAR ALGEBRA: PROOFS

In this appendix we prove the two quantitative statements from section 2, as well as a small collection of natural quantitative linear algebra statements which we will use freely throughout the paper.

We begin with a simple proposition concerning points bounded away from algebraic varieties.

Proposition B.1. *Let $I \subseteq \mathbb{R}[X_1, \dots, X_n]$ be an ideal with generators q_1, \dots, q_l , and suppose that $\mathbf{x} \in \mathbb{R}^n$ is a point with $\|\mathbf{x}\|_\infty \leq C$ and with $\text{dist}(\mathbf{x}, V(I)) \geq c$, for some absolute constants c and C . Then, there is some q_j such that $|q_j(\mathbf{x})| = \Omega_{c,C,I}(1)$.*

Proof. Suppose for contradiction that, for all $\varepsilon > 0$, there exists an $\mathbf{x} \in \mathbb{R}^n$ with $\|\mathbf{x}\|_\infty \leq C$ and with $\text{dist}(\mathbf{x}, V(I)) \geq c$, but with $|q_j(\mathbf{x})| < \varepsilon$ for every j . Taking a sequence of ε tending to 0, we get a corresponding sequence of \mathbf{x}_ε . Since all \mathbf{x}_ε lie in a compact set, there exists a convergent subsequence tending to some limit point \mathbf{x} . But then $\text{dist}(\mathbf{x}, V(I)) \geq c$ by the continuity of the dist function, and since $V(I)$ is closed in the Euclidean topology, yet $q_j(\mathbf{x}) = 0$ for every j , and so $\mathbf{x} \in V(I)$, a contradiction. \square

For this proposition it is easy to deduce the existence of rank matrices. We recall the statement from section 2.

Proposition B.2. *Suppose that $\|L\|_\infty \geq C$ and $\|L - L'\|_\infty \geq c$ for all matrices L' with $\text{rank}(L') < m$. Then L has a rank matrix M , with $|\det M| = \Omega_{c,C}(1)$. Furthermore, if \mathbf{v} is any vector in the row-space of L with coefficients bounded by C , then for $1 \leq i \leq m$ there exist coefficients a_i with $|a_i| = O_{c,C}(1)$ such that $\sum_{i=1}^m a_i \lambda_{ij} = v_j$ for all $1 \leq j \leq d$.*

Proof. For $k = \binom{d}{m}$, let q_1, \dots, q_k be the k polynomials on \mathbb{R}^{md} , a space which we identify with the space of m -by- d matrices, given by the k determinants of m -by- m submatrices. The variety V' , consisting of all matrices L' with rank less than m , is equal to $V(I)$, where $I \subseteq \mathbb{R}[X_1, \dots, X_{md}]$ is the ideal generated by the polynomials q_i . This is since row rank equals column rank, and linear independence of columns in a square matrix can be detected by the determinant. The assumption of the proposition is exactly that $\text{dist}(L, V') \geq c$.

Therefore we can apply Proposition B.1 to deduce that there must exist some $\delta > 0$, depending only on c and C , such that for all L with $\|L\|_\infty \geq C$ and $\text{dist}(L, V') \geq c$ there is some j with $|q_j(L)| \geq \delta$. The matrix whose determinant corresponds to the polynomial q_j is exactly the claimed rank matrix.

The statement about linear combinations of rows follows quickly from the existence of the rank matrix. Indeed, without loss of generality assume that the rank matrix is realised by columns 1 through m . The fact that the rows of L are linearly independent means that there are unique a_i such that $\sum_{i=1}^m a_i \lambda_{ij} = v_j$ for all $1 \leq j \leq d$. Restricting to $1 \leq j \leq m$, we observe that the a_i are forced to satisfy

$$\begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} = (M^T)^{-1} \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix},$$

and hence, by the properties of the rank matrix, $a_i = O_{c,C}(1)$ for all i . \square

We now consider the quantitative normal form algorithm. Let us recall the statement, and then give the proof.

Proposition B.3. *Let $(\psi_1, \dots, \psi_m) = \Psi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a system of real linear forms with coefficients bounded above in absolute value by C . Furthermore, suppose that there exists an absolute $c > 0$ such that for each i there exists some partition \mathcal{P}_i for which $\text{dist}(\Psi, V_{\mathcal{P}_i}) \geq c$. Then for each i there is an extension $\Psi' : \mathbb{R}^{n'} \rightarrow \mathbb{R}^m$ such that:*

- $n' \leq n + m - 1$.
- Ψ' is of the form

$$\Psi'(\mathbf{u}, w_1, \dots, w_{s+1}) := \Psi(\mathbf{u} + w_1 \mathbf{f}^1 + \dots + w_{s+1} \mathbf{f}^{s+1})$$

for some vectors $\mathbf{f}^k \in \mathbb{R}^n$ and some $s \geq 0$, such that $\|\mathbf{f}^k\|_\infty = O_{c,C}(1)$ for every k .

- Ψ' is in normal form with respect to ψ'_i .
- All the coefficients of ψ'_i attached to the variables w_k are $\Theta_{c,C}(1)$.

In the statement and proof, implied constants may depend on m and n .

Proof. Fix i , and let $\cup \mathcal{C}_k$ be the partition \mathcal{P}_i of $[m] \setminus \{i\}$ such that $\text{dist}(\Phi, V_{\mathcal{P}_i}) \geq c$. Suppose that \mathcal{P}_i has $s+1$ parts. Via Gaussian elimination, we may find vectors $\mathbf{f}^k \in \mathbb{R}^n$ which witness the fact that $\text{dist}(\Phi, V_{\mathcal{P}_i}) > 0$, i.e. for which $\psi_i(\mathbf{f}^k) = 1$ but $\psi_j(\mathbf{f}^k) = 0$ for all $j \in \mathcal{C}_k$. If given a free choice for one of the coordinates of \mathbf{f}^k , we set it to be 0.

We claim that the form

$$\Psi'(\mathbf{u}, w_1, \dots, w_{s+1}) := \Psi(\mathbf{u} + w_1 \mathbf{f}^1 + \dots + w_{s+1} \mathbf{f}^{s+1})$$

is a suitable extension of Ψ . Indeed, $\psi'_i(\mathbf{u}, w_1, \dots, w_{s+1})$ is the only one of the ψ'_j forms to use all of the w_k variables, and the coefficient of each w_k in ψ'_i is $\psi_i(\mathbf{f}^k)$, i.e. the w_k coefficient is 1. Also, $n' = n + s + 1$, which is at most $n + m - 1$.

We claim further that one may find¹⁶ such \mathbf{f}^k satisfying $\|\mathbf{f}^k\|_\infty \leq O_{c,C}(1)$, after which the proposition is proved. Indeed, consider a fixed k , and consider all the possible ways of implementing Gaussian elimination. When applying the algorithm to a fixed Ψ one decides a particular implementation based on which co-ordinates become zero or non-zero after one performs linear combinations of the rows, but for now consider the matrix on which the algorithm is applied to be an abstract unknown and consider all possible sequences of row combinations. Inspecting the algorithm, the co-ordinates of the claimed solution vector \mathbf{f}^k after a particular implementation are the evaluations of certain rational functions taken at the coefficients of Ψ . [We now identify Ψ with the coordinate vector in \mathbb{R}^{mn} of its coefficients.] It could be that Ψ is a pole of some of these functions, although we know that there is at least one implementation of the algorithm in which it is not.

Let Γ be the set of possible implementations of Gaussian elimination. The size $|\Gamma|$ is essentially $(1 + |\mathcal{C}_k|)!$, but for us it will be enough that $|\Gamma| = O(1)$. Now, for each $\gamma \in \Gamma$, let rational functions

$$\frac{p_{\gamma,1}(\Psi)}{q_{\gamma,1}(\Psi)}, \dots, \frac{p_{\gamma,n}(\Psi)}{q_{\gamma,n}(\Psi)}$$

be the n rational functions defining the claimed coefficients of \mathbf{f}^k . One may assume without loss of generality that, for all j , $p_{\gamma,j}, q_{\gamma,j} \in \mathbb{Z}[X_1, \dots, X_n]$ are co-prime polynomials, with coefficients of size $O(1)$. Now let

$$Q_\gamma := \text{lcm}_{j \leq n} q_{\gamma,j}.$$

Then it is clear that the affine algebraic variety $V(I)$, with I the ideal generated by the set of polynomials $\{Q_\gamma : \gamma \in \Gamma\}$, is contained in $V_{\mathcal{P}_i}$. Indeed, if $Q_\gamma(\Psi) = 0$ for all $\gamma \in \Gamma$ then there is no Gaussian elimination implementation which finds a solutions \mathbf{f}^k , which in turn implies that \mathcal{P}_i is not suitable for Ψ . In particular we have $\text{dist}(\Psi, V(I)) \geq c$.

¹⁶Even with $\text{dist}(\Psi, V_{\text{degen}}) \geq c$, there could be unbounded \mathbf{f}^k which satisfy $\psi_i(\mathbf{f}^k) = 1$ and $\psi_j(\mathbf{f}^k) = 0$ for all $j \in \mathcal{C}_k$.

Applying Proposition B.1 to the ideal I , we conclude that there is some $\gamma \in \Gamma$ such that $|Q_\gamma(\Psi)| = \Omega_{c,C}(1)$. In particular, we conclude that the solution vector $\mathbf{f}^{\mathbf{k}}$ obtained by the implementation γ has coefficients which are bounded by $O_{c,C}(1)$. \square

Let us illustrate the above proof with an example. Consider $n = 3$, $m = 2$, $i = 2$, and denote

$$\Psi = \begin{pmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \end{pmatrix}.$$

Then the partition \mathcal{P}_i consists of the singleton $\{1\}$, and implementing Gaussian elimination a certain way we have

$$\mathbf{f}^1 = \begin{pmatrix} s_{22}/(s_{11}s_{22} - s_{12}s_{21}) \\ -s_{21}/(s_{11}s_{22} - s_{12}s_{21}) \\ 0 \end{pmatrix}$$

as a solution, in the case where $s_{11}s_{22} - s_{12}s_{21}$ is non-zero. Of course if $s_{11}s_{23} - s_{13}s_{21}$ is non-zero too, we have another solution

$$\mathbf{f}^1 = \begin{pmatrix} s_{23}/(s_{11}s_{23} - s_{13}s_{21}) \\ 0 \\ -s_{21}/(s_{11}s_{23} - s_{13}s_{21}) \end{pmatrix}.$$

So, if one applied Gaussian elimination idly, one might end up with either of these two solutions. Unfortunately it could be the case that $\text{dist}(\Psi, V_{\mathcal{P}_i}) \geq c$ whilst one of these determinants, $s_{11}s_{22} - s_{12}s_{21}$ say, was non-zero yet $o(1)$. In this instance, applying the first implementation of the algorithm would not give a desirable solution vector \mathbf{f}^1 . Therefore we need some subtlety to ensure that we pick the correct implementation.

It is worth a brief discussion of why these quantitative subtleties do not arise in the setting of [10]. Indeed, assume that Ψ has rational coefficients of naive height at most C . Proceed with all the linear algebra from the previous lemma over \mathbb{Q} , and choose *any* implementation of Gaussian elimination which is valid for Ψ . As previously remarked, the co-ordinates of the solution vector $\mathbf{f}^{\mathbf{k}}$ are the evaluations of certain rational functions $\frac{p_i}{q_j}$ with $p_j, q_j \in \mathbb{Z}[X_1, \dots, X_n]$ co-prime, taken at the coefficients of Ψ . [We now once more identify Ψ with the coordinate vector in \mathbb{R}^{mn} of its coefficients.] By the construction of the algorithm,

$$\Psi \notin \bigcup_{j=1}^n \{\Psi' : q_j(\Psi') = 0\}.$$

The set on the right-hand side is closed in the Euclidean topology.

The key change from the situation over the reals is that there are only $O_C(1)$ many possible Ψ (since Ψ has rational coordinates of bounded height), and so with the above information we can immediately conclude that

$$\text{dist}(\Psi, \bigcup_{j=1}^n \{\Psi' : q_j(\Psi') = 0\}) \gg_C 1,$$

without needing to assume this as an extra hypothesis.

We continue this appendix by proving that the conditions on L in the statement of Theorem 1 imply that there exists a parametrisation Ψ of $\ker(L)$, introduced in equation (30), which admits a quantitative normal form extension. By the preceding proposition, it will be enough to show that there is such a parametrisation Ψ and for each i partitions \mathcal{P}_i with $\text{dist}(\Psi, V_{\mathcal{P}_i}) \geq c$. Yet by considering the partition in which every form ψ_k is in its own part, we see it is enough to find Ψ such that $\text{dist}(\Psi, V_{\text{degen}}) \gg 1$.

Another remark: one notices that in section 4 we in fact proceeded by parametrising the kernel of $S_{\text{rat}} + S_{\text{irrat}}$, rather than the kernel of L . However, it is easy to see that if L satisfies $\|L - L'\|_\infty \geq c$ for all matrices L' with $\text{rank}(L') < m$, $\|L\|_\infty \leq C$, and $\text{dist}(L, V_{\text{degen}}^*) \geq c$, then the same is true for both S and $S_{\text{rat}} + S_{\text{irrat}}$, albeit with the quantities c and C replaced by $\Omega_{c,C}(1)$ and $O_{c,C}(1)$ respectively. For notational clarity we prove everything to follow in terms of L , but the proofs apply identically to $S_{\text{rat}} + S_{\text{irrat}}$ (which is the required application in this paper).

We introduce the necessary ideas by proving a non-quantitative proposition, before extending to the full-strength result required.

Proposition B.4. *Suppose $\text{rank}(L) = m$ and $L \notin V_{\text{degen}}^*$. Then $K := \ker L$ has an orthonormal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_{d-m}\}$ such that the associated parametrisation $\Psi : \mathbb{R}^{d-m} \rightarrow K$ defined by $\Psi(\mathbf{u}) = \sum_{j=1}^{d-m} u_j \mathbf{v}_j$ satisfies $\Psi \notin V_{\text{degen}}$.*

Proof. We proceed by a simple duality argument. By the usual Gram-Schmidt process we may certainly find some orthonormal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_{d-m}\}$ for K , and suppose for contradiction that the associated Ψ satisfies $\Psi \in V_{\text{degen}}$. Then there are two forms ψ_i and ψ_j such that $\psi_i = \alpha \psi_j$ for some $\alpha \in \mathbb{R}$. Translating in terms of the basis, this implies that every basis vector \mathbf{v}_k is orthogonal to the vector $\alpha \mathbf{e}_i - \mathbf{e}_j$, and hence all of K is orthogonal to $\alpha \mathbf{e}_i - \mathbf{e}_j$.

Viewing the row space of L as an m -dimensional subspace of \mathbb{R}^d , K is exactly the orthogonal complement to this space. In particular the row space of L exactly consists of all those vectors which are orthogonal to K . So we conclude that $\alpha \mathbf{e}_i - \mathbf{e}_j$ is in the row space of L , i.e. there exists a non-zero vector in the row space of L with two or fewer non-zero coordinates. So, by definition, $L \in V_{\text{degen}}^*$. \square

Proposition B.5. *Suppose $\|L - L'\|_\infty \geq c$ for all matrices L' with $\text{rank}(L') < m$, $\|L\|_\infty \leq C$ and $\text{dist}(L, V_{\text{degen}}^*) \geq c$. Then $K := \ker L$ has an orthonormal basis $\{\mathbf{v}_1, \dots, \mathbf{v}_{d-m}\}$ such that the associated parametrisation $\Psi : \mathbb{R}^{d-m} \rightarrow K$ defined by $\Psi(\mathbf{u}) = \sum_{j=1}^{d-m} u_j \mathbf{v}_j$ satisfies $\|\Psi\|_\infty = O_{c,C}(1)$ and $\text{dist}(\Psi, V_{\text{degen}}) = \Omega_{c,C}(1)$.*

Proof. Pick $\{\mathbf{v}_1, \dots, \mathbf{v}_{d-m}\}$ any orthonormal basis for K , and let Ψ be the associated parametrisation. Certainly $\|\Psi\|_\infty = O_{c,C}(1)$ since indeed $\|\Psi\|_\infty \leq 1$, as the chosen basis is orthonormal, but suppose for contradiction that $\text{dist}(\Psi, V_{\text{degen}}) \leq \eta$ for some very small parameter η . In other words, without loss of generality, there exist vectors $\{\mathbf{w}_1, \dots, \mathbf{w}_{d-m}\}$ and $\alpha \in \mathbb{R}$ such that for all $j \in [d-m]$ we have $\|\mathbf{w}_j - \mathbf{v}_j\|_\infty \leq \eta$ and $\mathbf{w}_j \cdot (\alpha \mathbf{e}_1 - \mathbf{e}_2) = 0$.

We claim that $L = L' + E$, where $\|E\|_\infty = O_C(\eta)$ and the vector $\alpha \mathbf{e}_1 - \mathbf{e}_2$ is in the row space of L' , whereupon we can immediately conclude the proof, since then $\text{dist}(L, V_{\text{degen}}^*) \leq O_C(\eta)$, which is a contradiction if η is small enough.

Indeed, for $i \in [m]$ let $\boldsymbol{\lambda}_i \in \mathbb{R}^d$ denote the i^{th} row of L . Then for all $i \in [m]$, $j \in [d-m]$, we have $|\boldsymbol{\lambda}_i \cdot \mathbf{w}_j| = |\boldsymbol{\lambda}_i \cdot (\mathbf{w}_j - \mathbf{v}_j)| = O_C(\eta)$. Letting $\boldsymbol{\lambda}_i'$ denote the projection of $\boldsymbol{\lambda}_i$ onto $\text{span}(\mathbf{w}_j : j \in [d-m])^\perp$, we claim that this observation implies that $\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_i'\|_\infty = O_C(\eta)$. Indeed, for each j we have $\|\mathbf{w}_j\|_\infty = \Theta(1)$ provided that η is small enough. Further, for $i \neq j$ note that

$$\begin{aligned} |\mathbf{w}_i \cdot \mathbf{w}_j| &\leq |\mathbf{w}_i \cdot (\mathbf{w}_j - \mathbf{v}_j)| + |\mathbf{v}_i \cdot \mathbf{v}_j| + |(\mathbf{w}_i - \mathbf{v}_i) \cdot \mathbf{v}_j| \\ &= |\mathbf{w}_i \cdot (\mathbf{w}_j - \mathbf{v}_j)| + |(\mathbf{w}_i - \mathbf{v}_i) \cdot \mathbf{v}_j| \\ &= O(\eta), \end{aligned}$$

so for η small enough we certainly have the \mathbf{w}_j being linear independent. Further, extending to a full basis $\mathcal{B} := \{\mathbf{w}_1, \dots, \mathbf{w}_d\}$ of \mathbb{R}^d by choosing orthogonal vectors of unit length, the change of basis matrix between the standard basis and this basis has determinant $\Theta(1)$. Expanding $\boldsymbol{\lambda}_i$ in terms of the basis \mathcal{B} , and using the fact that $|\boldsymbol{\lambda}_i \cdot \mathbf{w}_j| = O_C(\eta)$ for all $j \in [d - m]$, we have $\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_i'\|_{2,\mathcal{B}} = O_C(\eta)$, where $\|\cdot\|_{2,\mathcal{B}}$ denotes the L^2 norm with respect to the basis \mathcal{B} . Therefore $\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_i'\|_\infty = O_C(\eta)$ as claimed.

Finally, let L' be the m -by- d matrix with rows given by the $\boldsymbol{\lambda}_i'$. Since for η small enough the $\boldsymbol{\lambda}_i'$ are linearly independent, it follows by dimension counting that the row space of L' is the entirety of $\text{span}(\mathbf{w}_j : j \in [d - m])^\perp$. In particular the row space of L' contains the vector $\alpha \mathbf{e}_1 - \mathbf{e}_2$, and the proposition is proved. \square

To conclude this appendix, we adapt these ideas to show

Proposition B.6. *Suppose $\|L - L'\|_\infty \geq c$ for all matrices L' with $\text{rank}(L') < m$, $\|L\|_\infty \leq C$ and $\text{dist}(L, V_{\text{degen}}^*) \geq c$. Let $K := \ker(L)$. Then, if $\mathbf{v} \in K^\perp$, we have $L\mathbf{v} \in B_\varepsilon^m$ only when $\mathbf{v} \in R$, where $R \subseteq [-O_{c,C,\varepsilon}(1), O_{c,C,\varepsilon}(1)]^d$ is some convex region.*

Proof. Suppose $L\mathbf{v} \in B_\varepsilon^m$. In the notation of the previous proof, we have $|\boldsymbol{\lambda}_i \cdot \mathbf{v}| \leq \varepsilon$ for all $i \in [m]$. Since $\text{dist}(L, V_{\text{degen}}^*) \geq c$, we certainly have that the $\boldsymbol{\lambda}_i$ are linearly independent. Moreover, it follows from the second conclusion of Proposition B.2 that we may extend the set $\{\boldsymbol{\lambda}_i : i \in [m]\}$ by orthogonal vectors of unit length to form a basis $\{\boldsymbol{\lambda}_i : i \in [d]\}$ for \mathbb{R}^d such that for all $k, j \leq d$ we have

$$\sum_{i=1}^d a_{ki} \lambda_{ij} = \delta_{kj}$$

for coefficients a_{ki} satisfying $|a_{ki}| = O_{c,C}(1)$. Indeed, fix k , and note that the standard basis vector \mathbf{e}_k is equal to $\mathbf{x}_k + \mathbf{y}_k$, where $\mathbf{x}_k \in \text{span}(\boldsymbol{\lambda}_i : i \in [m])$ and $\mathbf{y}_k \in \text{span}(\boldsymbol{\lambda}_i : m+1 \leq i \leq d)$. These two vectors are orthogonal by construction, so in particular $\|\mathbf{x}_k\|_2^2 + \|\mathbf{y}_k\|_2^2 = 1$, and hence $\|\mathbf{x}_k\|_\infty, \|\mathbf{y}_k\|_\infty \ll 1$. By Proposition B.2 applied to \mathbf{x}_k we get $|a_{ki}| = O_{c,C}(1)$ for $i \leq m$, and the orthonormality of $\{\boldsymbol{\lambda}_i : m+1 \leq i \leq d\}$ implies that $|a_{ki}| = O(1)$ for $m+1 \leq i \leq d$.

Now notice that $\text{span}(\boldsymbol{\lambda}_i : m+1 \leq i \leq d)$ is exactly equal to K . Letting L' be the d -by- d matrix whose rows are $\boldsymbol{\lambda}_i$, and using the fact that $\mathbf{v} \in K^\perp$, we have that $L'\mathbf{v} = \mathbf{w}$ for some vector \mathbf{w} with $\|\mathbf{w}\|_\infty \leq \varepsilon$. Premultiplying by the matrix $A = (a_{ki})$, we immediately get $\mathbf{v} = A\mathbf{w}$, and hence $\|\mathbf{v}\|_\infty = O_{c,C,\varepsilon}(1)$. The region R is therefore bounded, and is clearly convex, and so the proposition is proved. \square

APPENDIX C. EQUIDISTRIBUTION

In this appendix we prove two propositions required in section 3, and rigorously define the metric and the measure we are using on tori $T \leq \mathbb{T}^m$.

The metric $\|\mathbf{x} - \mathbf{y}\|_T$ for $\mathbf{x}, \mathbf{y} \in T$ will simply be the restriction of the usual metric $\|\mathbf{x} - \mathbf{y}\|_{\mathbb{T}^m}$. To define the measure, let $J \in SL_m(\mathbb{Z})$ be any matrix such that $J(T) = \mathbb{T}^{\dim(T)} \times \{0\}^{m-\dim(T)}$. Then one can define the measure μ_T on T as the pull-back¹⁷ $J^*\mu$ of Lebesgue measure on $\mathbb{T}^{\dim(T)}$, i.e. on $[0, 1]^{\dim(T)}$. In other words, $\mu_T(U) := \mu(J(U))$ for open sets U and then extend uniquely. This definition is manifestly independent of the choice of J : if J_1 and J_2 are two choices, then $J_1 J_2^{-1} : \mathbb{T}^{\dim(T)} \rightarrow \mathbb{T}^{\dim(T)}$ is measure preserving since it is a linear map with determinant 1.

Now let us state and prove the two propositions required in section 3.

¹⁷According to taste one could define the measure as a pushforward of μ under J^{-1} .

Proposition C.1. *Let T_1, \dots, T_l be sub-tori of \mathbb{T}^m , each with complexity at most A_j respectively, and let $A = \max(A_j : j \in [l])$. Then the torus $T := T_1 + \dots + T_l$ has complexity at most $O_{l,m}(A^{O_{l,m}(1)})$.*

Proof. This follows easily from simple consideration of integer lattices. Indeed, without loss of generality we may assume that all the tori are non-trivial, and it is also clearly enough to assume that $l = 2$, as the general claim follows by induction. Now, choose $P_1, P_2 \in SL_m(\mathbb{Z})$ such that $P_j(T_j) = \mathbb{T}^{m_j} \times \{0\}^{m-m_j}$ for $j = 1, 2$. By assumption, one may choose $\|P_j\|_\infty \leq A_j$ for $j = 1, 2$. Now, for $i = 1, \dots, m_j$, let vector $\mathbf{v}_i^j := P_j^{-1}(\mathbf{e}_i)$. Also, let $V_j = P_j^{-1}(\mathbb{R}^{m_j} \times \{0\}^{m-m_j})$. Then the vectors \mathbf{v}_i^j are in \mathbb{Z}^m , satisfy $\|\mathbf{v}_i^j\|_\infty \leq A_j^{O_m(1)}$, and form a \mathbb{Z} -basis for the lattice $\mathbb{Z}^m \cap V_j$. We call this lattice Λ_j .

Now consider the set of vectors $S = \{\mathbf{v}_1^1, \dots, \mathbf{v}_{m_1}^1, \mathbf{v}_1^2, \dots, \mathbf{v}_{m_2}^2\}$. These \mathbb{Z} -span the lattice $\Lambda_1 + \Lambda_2$. This implies that one can find a \mathbb{Z} -basis $\{\mathbf{w}_1, \dots, \mathbf{w}_d\}$ for $\Lambda_1 + \Lambda_2$ where one has $\|\mathbf{w}_i\|_\infty \ll_m A^{O_m(1)}$ for every i .

Indeed, we prove the more general statement that any lattice $\Lambda \leq \mathbb{Z}^m$ which has a \mathbb{Q} -basis $\{\mathbf{v}_1, \dots, \mathbf{v}_d\}$ of integer vectors with $\|v_i\|_\infty \leq A$ for all i also possesses a \mathbb{Z} -basis $\{\mathbf{w}_1, \dots, \mathbf{w}_d\}$ with $\|\mathbf{w}_i\|_\infty \ll_m A$ for all i . We proceed by induction on the dimension d of Λ . If $d = 1$, then let \mathbf{w}_1 be the vector in $\Lambda \setminus \{0\}$ such that $\|\mathbf{w}_1\|_2$ is minimal. Then $\{\mathbf{w}_1\}$ \mathbb{Z} -spans Λ , and $\|\mathbf{w}_1\|_2 \leq \|\mathbf{v}_1\|_2 \ll_m A$.

Now assume $d \geq 2$. Letting $\Lambda' := \Lambda \cap \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_{d-1})$, for the induction step we may assume without loss of generality that the set $\{\mathbf{v}_1, \dots, \mathbf{v}_{d-1}\}$ is a \mathbb{Z} -basis of Λ' . Then, $\{\mathbf{v}_1, \dots, \mathbf{v}_{d-1}, \mathbf{v}_d\}$ is certainly a \mathbb{Z} -basis for some sublattice $\Lambda_{\text{sub}} \leq \Lambda$. Now, let \mathbf{w}_d be a vector in $\Lambda \setminus \text{span}_{\mathbb{R}}(\mathbf{v}_1, \dots, \mathbf{v}_{d-1})$ such that $\text{dist}_{\text{Eucl}}(\mathbf{w}_d, \text{span}_{\mathbb{R}}(\mathbf{v}_1, \dots, \mathbf{v}_{d-1}))$ is minimal. For any such vector, $\{\mathbf{v}_1, \dots, \mathbf{v}_{d-1}, \mathbf{w}_d\}$ is a \mathbb{Z} -basis of Λ , as there are no elements of Λ within the parallelepiped which these vectors form. The norm $\|\mathbf{w}_d\|_2$ may be very large but, by subtracting off suitable elements of Λ' from such a \mathbf{w}_d , we may without loss of generality assume that \mathbf{w}_d is contained within the fundamental parallelepiped of Λ_{sub} formed by $\{\mathbf{v}_1, \dots, \mathbf{v}_{d-1}, \mathbf{v}_d\}$. Every vector in this parallelepiped has Euclidean norm at most $O_m(A)$, and hence $\|\mathbf{w}_d\|_2 \ll_m A$. Therefore $\|\mathbf{w}_d\|_\infty \ll_m A$ and the claim is proved.

So we have a \mathbb{Z} -basis $\{\mathbf{w}_1, \dots, \mathbf{w}_d\}$ for $\Lambda_1 + \Lambda_2$ with $\|\mathbf{w}_i\|_\infty \ll_m A^{O_m(1)}$ for every i . But the existence of such a \mathbb{Z} -basis for $\Lambda_1 + \Lambda_2$ implies that $T_1 + T_2$ has complexity at most $O_m(A^{O_m(1)})$ as required. Indeed, it is immediately seen that $\Lambda_1 + \Lambda_2$ is equal to $\mathbb{Z}^m \cap (V_1 + V_2)$, and therefore no-integer points are found within the fundamental parallelepiped formed by $\{\mathbf{w}_1, \dots, \mathbf{w}_d\}$. By following the same procedure as above, one may extend $\{\mathbf{w}_1, \dots, \mathbf{w}_d\}$ to a \mathbb{Z} -basis $\{\mathbf{w}_1, \dots, \mathbf{w}_m\}$ for the lattice \mathbb{Z}^m such that $\|\mathbf{w}_i\|_\infty \ll_m A^{O_m(1)}$ for every i . But then the m -by- m matrix P whose columns are given by the vectors \mathbf{w}_i defines an isomorphism from \mathbb{Z}^m to \mathbb{Z}^m . In particular both P and P^{-1} have integer entries, which implies¹⁸ that $\det(P) = \pm 1$. By choosing \mathbf{w}_d of the appropriate sign, we may ensure that $\det(P) = 1$, i.e. $P \in SL_m(\mathbb{Z})$.

Finally, the coefficients of P^{-1} have size at most $O_m(A^{O_m(1)})$, and P^{-1} maps $T_1 + T_2$ onto $\mathbb{T}^d \times \{0\}^{m-d}$. Therefore the complexity of $T_1 + T_2$ is bounded as desired. \square

Proposition C.2. *Let C_1 be a constant, and let $\theta_1, \dots, \theta_l \in \mathbb{R}^m$ each be $(\mathcal{F}(A_j), N)$ -irrational in sub-tori T_1, \dots, T_l of \mathbb{T}^m of complexity at most A_j respectively, where \mathcal{F} is a sufficiently quickly growing growth function. Then for every Lipschitz function*

¹⁸Immediately from the consideration $1 = \det(PP^{-1}) = \det(P) \det(P^{-1})$.

$F : T \longrightarrow \mathbb{C}$,

$$\left| \frac{1}{N^l} \sum_{\substack{n_1, \dots, n_l \\ n_j \leq N \forall j}} F \left(\sum_{j=1}^l \theta_j n_j \right) - \int_{\mathbf{x} \in T} F(\mathbf{x}) d\mu_T(\mathbf{x}) \right| \ll_{m,l} A^{-C_1} \|F\|_{\text{Lip}},$$

where $A = \max(A_j : j \in [l])$

Proof. This is an easy adaptation of the proof of [13, Theorem 3.1], which we stated in this paper as Theorem 18. We sketch the required alterations.

Suppose first that $T = \mathbb{T}^t \times \{0\}^{m-t}$, for some dimension $t \geq 1$. Then if the conclusion to the theorem is false, one may proceed identically to the proof of [13, Theorem 3.1], taking $\delta = A^{-C_1}$. One establishes that there is some $\mathbf{k} \in \mathbb{Z}^t$ with $0 < \|\mathbf{k}\|_1 < A^{-O_t(C_1)}$ such that

$$\left| \frac{1}{N^l} \sum_{n_1, \dots, n_l \leq N} e(\mathbf{k} \cdot \sum_{j=1}^l \theta_j n_j) \right| \gg_{t,l} A^{-O_t(C_1)}.$$

The left-hand side may be recast as a product, so we have

$$\prod_{j=1}^l \left| \frac{1}{N} \sum_{n \leq N} e(n\mathbf{k} \cdot \theta_j) \right| \gg_{t,l} A^{-O_t(C_1)}. \quad (55)$$

Note that if \mathbf{k} were in the annihilator of every T_j , i.e. $\mathbf{k} \cdot \theta_j \in \mathbb{Z}$ for every $\theta_j \in T_j$, then \mathbf{k} would be in the annihilator of T itself. But $T = \mathbb{T}^t$, and therefore \mathbf{k} would be $\mathbf{0}$, a contradiction. Therefore, there exists some $j \leq l$ such that \mathbf{k} is not in the annihilator of T_j , and this observation will be enough to contradict (55).

Indeed, let $J_j \in SL_t(\mathbb{Z})$ be some given matrix with $\|J_j\|_\infty \leq A$ such that $J_j(T_j) = \mathbb{T}^{\dim(T_j)} \times \{0\}^{t-\dim(T_j)}$. Certainly we have $(J_j^T)^{-1}\mathbf{k} \in \mathbb{Z}^t$, with $0 < \|(J_j^T)^{-1}\mathbf{k}\|_1 \leq A^{O_t(1)} < \mathcal{F}(A)$. Therefore, since θ_j is $(\mathcal{F}(A), N)$ irrational in T_j ,

$$\left| \sum_{n \leq N} e(\mathbf{k} \cdot \theta_j) \right| \ll \min \left(1, \frac{1}{N \|\mathbf{k} \cdot \theta_j\|_{\mathbb{T}}} \right) \ll \min \left(1, \frac{1}{N \|(J_j^T)^{-1}\mathbf{k} \cdot J_j(\theta_j)\|_{\mathbb{T}}} \right) \ll \mathcal{F}(A)^{-1}.$$

This directly contradicts (55), provided $\mathcal{F}(A) \geq A^{C_2}$ for some large constant C_2 .

We now reduce the case of general T to the above. Indeed, let $J \in SL_m(\mathbb{Z})$ be the chosen matrix with $\|J\|_\infty \leq A^{O_m(1)}$ and $J(T) = \mathbb{T}^{\dim(T)} \times \{0\}^{m-\dim(T)}$. For $F : T \longrightarrow \mathbb{C}$ bounded and Lipschitz, we have

$$\begin{aligned} & \left| \frac{1}{N^l} \sum_{n_1, \dots, n_l \leq N} F \left(\sum_{j=1}^l \theta_j n_j \right) - \int_{\mathbf{x} \in T} F(\mathbf{x}) d\mu_T(\mathbf{x}) \right| \\ &= \left| \frac{1}{N^l} \sum_{n_1, \dots, n_l \leq N} (FJ^{-1}) \left(\sum_{j=1}^l J(\theta_j) n_j \right) - \int_{\mathbf{x} \in [0,1]^{\dim(T)} \times \{0\}^{m-\dim(T)}} (FJ^{-1})(\mathbf{x}) d\mathbf{x} \right| \\ &\leq_{l,m} A^{-C_1} \|FJ^{-1}\|_{\text{Lip}} \end{aligned} \quad (56)$$

for some large constant C_1 depending on the choice of \mathcal{F} , by applying the previous argument to function FJ^{-1} and the torus $\mathbb{T}^{\dim(T)} \times \{0\}^{m-\dim(T)}$. Note, by considering the composition $J_j J^{-1}$, that indeed the tori $J(T_j)$ have complexity at most $A^{O_t(1)}$. The

proposition is concluded once we have observed that, for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$,

$$\begin{aligned} \|(FJ^{-1})(\mathbf{u}) - (FJ^{-1})(\mathbf{v})\|_{\mathbb{T}^m} &\leq \|F\|_{\text{Lip}} \|J^{-1}\mathbf{u} - J^{-1}\mathbf{v}\|_{\mathbb{T}^m} \\ &= \|F\|_{\text{Lip}} \|J^{-1}(\mathbf{u}' - \mathbf{v}')\|_{\mathbb{T}^m} \\ &\leq \|F\|_{\text{Lip}} \|J^{-1}(\mathbf{u}' - \mathbf{v}')\|_{\infty} \\ &\ll \|F\|_{\text{Lip}} \|J^{-1}\|_{op} \|\mathbf{u}' - \mathbf{v}'\|_{\infty} \\ &= \|F\|_{\text{Lip}} \|J^{-1}\|_{op} \|\mathbf{u} - \mathbf{v}\|_{\mathbb{T}^m}, \end{aligned}$$

where \mathbf{u}' and \mathbf{v}' are coset representatives of $\mathbf{u} \bmod \mathbb{Z}^m$ and $\mathbf{v} \bmod \mathbb{Z}^m$ respectively, chosen so that $\|\mathbf{u}' - \mathbf{v}'\|_{\infty} = \|\mathbf{u} - \mathbf{v}\|_{\mathbb{T}^m}$. (Note that J^{-1} has integer entries, so the above manipulation is valid.) Since $\|J\|_{\infty} \leq A$ and $\det J = 1$ we have $\|J^{-1}\|_{op} \leq A^{O_m(1)}$. Therefore $\|FJ^{-1}\|_{\text{Lip}} \leq \|F\|_{\text{Lip}} A^{O_m(1)}$, and this proves the proposition. \square

APPENDIX D. LIPSCHITZ FUNCTIONS

Here we collect two properties of Lipschitz functions required in the paper. As can be seen in light of the equidistribution definitions, it is cleaner to consider equidistribution with respect to Lipschitz (or at the very least continuous) functions, rather than sharp cut-offs. However, since we are ultimately interested in such sharp cut-offs, such approximation results can be useful.

Lemma D.1. [10, Corollary A.3] *Let $K \subset [-N, N]^m$ be a convex set, and $\sigma \in (0, 1)$. Then there exist Lipschitz functions $F_{\sigma}, G_{\sigma} : \mathbb{R}^m \rightarrow [0, 1]$ supported on $[-2N, 2N]^m$, both with Lipschitz constant at most $O(\frac{1}{\sigma N})$, such that $1_K = F_{\sigma} + O(G_{\sigma})$ and $\int_{\mathbb{R}^m} G_{\sigma}(\mathbf{x}) d\mathbf{x} = O_m(\sigma N^m)$. Furthermore, $F(\mathbf{x}) \geq 1_K(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^m$.*

This lemma allows us to combine the analytic theory of equidistribution discussed above with a variety of soft metric geometry observations.

If $K \subset [-\frac{1}{10}, \frac{1}{10}]^m$ is considered as a subset of the torus \mathbb{T}^m , then the Lipschitz functions in the above lemma may clearly also be considered as Lipschitz functions on \mathbb{T}^m . We will need a result for other tori T .

Lemma D.2. *Let $T \leq \mathbb{T}^m$ be a torus of complexity at most A . Let $K \subset \mathbb{R}^m$ be a small convex set of diameter $O(A^{-C})$, for some large enough constant C , and suppose further that $K \bmod \mathbb{Z}^m \subset T$. Let $\sigma \in (0, 1)$. Then there exist Lipschitz functions $F_{\sigma}, G_{\sigma} : T \rightarrow [0, 1]$, both with Lipschitz constant at most $O(A^{O_m(1)} \frac{1}{\sigma})$, such that $1_{K \bmod \mathbb{Z}^m} = F_{\sigma} + O(G_{\sigma})$ and $\int_T G_{\sigma}(\mathbf{x}) d\mu_T(\mathbf{x}) = O_m(\sigma)$. Furthermore, $F(\mathbf{x}) \geq 1_{K \bmod \mathbb{Z}^m}(\mathbf{x})$ for all $\mathbf{x} \in T$.*

Proof. One argues very similarly to the final section of the proof of Proposition C.2. Let $J \in SL_m(\mathbb{Z})$ be as before. The assumption on the diameter of K implies that the image $J(K) \in \mathbb{T}^{\dim T}$ is a small connected convex domain. Then we construct Lipschitz functions $F_{\sigma}^*, G_{\sigma}^* : \mathbb{T}^{\dim T} \rightarrow [0, 1]$ as in Lemma D.1. Then finally we define $F_{\sigma}(\mathbf{x}) := F_{\sigma}^*(J\mathbf{x})$ and $G_{\sigma}(\mathbf{x}) := G_{\sigma}^*(J\mathbf{x})$.

By the definition of μ_T we have

$$\int_T G_{\sigma}(\mathbf{x}) d\mu_T(\mathbf{x}) = \int_{\mathbb{T}^{\dim T}} G_{\sigma}^*(\mathbf{x}) d\mu(\mathbf{x}) = O_m(\sigma),$$

and by the argument from Proposition C.2 (with J^{-1} replaced by J), we have

$$\|F_{\sigma}\|_{\text{Lip}}, \|G_{\sigma}\|_{\text{Lip}} = O(A^{O_m(1)} \sigma).$$

\square

Thirdly, we give a statement concerning the approximation of Lipschitz functions by short exponential sums.

Lemma D.3. *Let $X > 2$, and let $F : \mathbb{R}^d \rightarrow \mathbb{C}$ be a Lipschitz function, supported on $[-O(N), O(N)]^d$, such that $\|F\|_\infty \leq 1$ and $\|F\|_{\text{Lip}} \leq M$.*

$$F_\sigma(\mathbf{x}) = \int_{\|\boldsymbol{\xi}\|_\infty \leq X} c_X(\boldsymbol{\xi}) e(\boldsymbol{\xi} \cdot \mathbf{x}) d\boldsymbol{\xi} + O\left(NM \frac{\log X}{X}\right) \quad (57)$$

for some function $c_X(\boldsymbol{\xi})$ with $\|c_X(\boldsymbol{\xi})\|_\infty \leq 1$.

This lemma is very similar to [9, Lemma A.9], and may be easily proved by adapting that standard harmonic analysis argument from \mathbb{T}^d to \mathbb{R}^d . Another similar device may be found in the proof of [13, Theorem 3.1].

REFERENCES

- [1] A. Baker. On some diophantine inequalities involving primes. *J. Reine Angew. Math.*, 228:166–181, 1967.
- [2] R. C. Baker. *Diophantine inequalities*, volume 1 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1986. Oxford Science Publications.
- [3] H. Davenport. *Analytic methods for Diophantine equations and Diophantine inequalities*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 2005. With a foreword by R. C. Vaughan, D. R. Heath-Brown and D. E. Freeman, Edited and prepared for publication by T. D. Browning.
- [4] H. Davenport and H. Heilbronn. On indefinite quadratic forms in five variables. *J. London Math. Soc.*, 21:185–193, 1946.
- [5] Sean Eberhard. The abelian arithmetic regularity lemma. Unpublished note, arXiv:1606.09303.
- [6] D. Eric Freeman. Asymptotic lower bounds and formulas for Diophantine inequalities. In *Number theory for the millennium, II (Urbana, IL, 2000)*, pages 57–74. A K Peters, Natick, MA, 2002.
- [7] W. T. Gowers. A new proof of Szemerédi’s theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001.
- [8] W. T. Gowers and J. Wolf. The true complexity of a system of linear equations. *Proc. Lond. Math. Soc.* (3), 100(1):155–176, 2010.
- [9] Ben Green and Terence Tao. Quadratic uniformity of the Möbius function. *Ann. Inst. Fourier (Grenoble)*, 58(6):1863–1935, 2008.
- [10] Ben Green and Terence Tao. Linear equations in primes. *Ann. of Math.* (2), 171(3):1753–1850, 2010.
- [11] Ben Green and Terence Tao. The Möbius function is strongly orthogonal to nilsequences. *Ann. of Math.* (2), 175(2):541–566, 2012.
- [12] Ben Green and Terence Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. *Ann. of Math.* (2), 175(2):465–540, 2012.
- [13] Ben Green and Terence Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. *Ann. of Math.* (2), 175(2):465–540, 2012.
- [14] Ben Green, Terence Tao, and Tamar Ziegler. An inverse theorem for the Gowers $U^{s+1}[N]$ -norm. *Ann. of Math.* (2), 176(2):1231–1372, 2012.
- [15] G. A. Margulis. Discrete subgroups and ergodic theory. In *Number theory, trace formulas and discrete groups (Oslo, 1987)*, pages 377–398. Academic Press, Boston, MA, 1989.
- [16] Randall McCutcheon. FVIP systems and multiple recurrence. *Israel J. Math.*, 146:157–188, 2005.
- [17] Wolfgang Müller. Systems of quadratic Diophantine inequalities. *J. Théor. Nombres Bordeaux*, 17(1):217–236, 2005.
- [18] Scott T. Parsell. Irrational linear forms in prime variables. *J. Number Theory*, 97(1):144–156, 2002.
- [19] Scott T. Parsell. On simultaneous diagonal inequalities. III. *Q. J. Math.*, 53(3):347–363, 2002.
- [20] T. Tao. An inverse theorem for the continuous gowers uniformity norm. Blog post, available at <https://terrytao.wordpress.com/2015/07/22/an-inverse-theorem-for-the-continuous-gowers-uniformity-norm/>.
- [21] Terence Tao. *Higher order Fourier analysis*, volume 142 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.
- [22] Terence Tao and Van H. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. Paperback edition [of MR2289012].
- [23] R. C. Vaughan. *The Hardy-Littlewood method*, volume 125 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, second edition, 1997.

ALED WALKER, MATHEMATICAL INSTITUTE, UNIVERSITY OF OXFORD, ANDREW WILES BUILDING,
RADCLIFFE OBSERVATORY QUARTER, WOODSTOCK ROAD, OXFORD, OX2 6GG.
`walker@maths.ox.ac.uk`